



MADURAI KAMARAJ UNIVERSITY

(University with Potential for Excellence)

DISTANCE EDUCATION



B.Sc., (Mathematics)

THIRD YEAR

**MODERN ALGEBRA AND
LINEAR ALGEBRA**

PAPER - VIII

Recognized By DEC
www.mkudde.org

\$ 89



MADURAI KAMARAJ UNIVERSITY

(University with Potential for Excellence)

DISTANCE EDUCATION



B.Sc. (MATHEMATICS)

Third Year

UNIT 1 - 10

PAPER - VIII

MODERN ALGEBRA AND LINEAR ALGEBRA

Recognized by DEC

www.mkudde.org

1000 COPIES



MADURAI KAMARAJ UNIVERSITY

(University with Potential for Excellence)

DISTANCE EDUCATION



B.Sc. (MATHEMATICS)

Third Year

UNIT I - 10

PAPER - VIII

MODERN ALGEBRA AND LINEAR ALGEBRA

Group properties, Commutative groups, Ideals

Group Lagrange's theorem and its converse

Recognized by DEO

www.mkudde.org

JEGN/1000 COPIES

Third Year

Dear Student,

We welcome you as a Student of the Third year B.sc degree Course in mathematics.

This paper – VIII deals with Modern Algebra and Linear Algebra. The learning material for this paper will be supplemented by Contact seminars.

Learning through the Distance Education mode, as you are all aware, involves self – learning and self – assessment and in this regard you are expected to put in disciplined and dedicated effort.

On our part, we assume of our guidance and support.

With best wishes.

SYLLABUS

B.Sc., Mathematics – Third year –

Major Paper – VIII

Modern Algebra and Linear Algebra

UNIT – 1:

Relation – Types of Relations – Binary Operations – Peano's

Postulates – principle of induction – Simple problems – Law of Trichotomy.

UNIT – 2:

Subgroups - Definitions and examples - Centre - Normalizer -

Intersection and Union of Subgroups - Permutations - Cycles

and Transpositions - Even and odd permutations - S_n and A_n -

Cycle groups - Definition and examples - Cyclic groups are

abelian - A group is cycle groups are cyclic - Theorem on the

number of generator of cyclic groups.

UNIT – 3:

Costs and their properties – Congruence relation modulo a

subgroup Lagrange's theorem and its consequences – The order

of an element of a finite group divides the order of the group. A

group of prime order is cyclic – A group has no proper

subgroup if it is cyclic group of prime order – Euler's theorem

– Fermat's theorem.

Normal subgroups – Equivalent conditions for a subgroup to be

normal – Any subgroup of an abelian group is normal – A

subgroup of index 2 is normal – Intersection of two normal

subgroups – Intersection of a subgroup and a normal subgroup.

If a subgroup has exactly one subgroup of given order then it is

normal – Quotient group.

UNIT – 4:

Homomorphism – Types of homomorphisms – Homomorphism

with reference to identity – Inverse and order of an element – its

properties – Kernel of homomorphism – Homomorphic image of

an abelian group is abelian – and that of a cyclic group is cyclic

– Isomorphism – Isomorphism is an equivalence relation among

groups – Any infinite cyclic group is isomorphic to the group of

integers – Any finite cyclic group of order n is isomorphic to

Z_n .

UNIT – 5:

Cayley's theorem – The fundamental theorem of homomorphism –

Rings - Definitions and examples – Elementary properties.

UNIT – 6:

Vectorspaces – Definition and Examples – Subspaces – Linear Transformations – Fundamental theorem of Homomorphism.

UNIT – 7:

Span of a set – Linear Independence – Basis and Dimension – Rank and Nullity Matrix and Linear Transformation.

UNIT – 8:

Inner product space – Definition and examples – Orthogonality – orthogonal complement.

UNIT – 9:

Matrices – Elementary transformation – inverse – Rank – Test for consistency – Solving linear equations – Cayley's Hamilton's theorem – Eigen values and Eigen vectors.

UNIT – 10:

Bilinear forms – Matrix of a bilinear form – Quadratic forms – Reduction to quadratic forms.

Text Book:

1. Modern Algebra by Dr. S. Arumugam

Publications: Sci Tech, 2006.

CONTENTS

UNIT NO.	SCHEME OF LESSONS	PAGE NO.
Unit – 1	RELATIONS	1
1.1	Introduction	1
1.2	Types of Relations	2
1.3	Binary operations	19
1.4	Peano's Postulates	25
1.5	Law of Trichotomy	27
1.6	Principle of Induction	32
Unit – 2	SUBGROUPS	41
2.0	Introduction	41
2.1	Subgroups	52
2.2	Permutation Groups	62
2.3	Cyclic Groups	73
Unit – 3	COSETS AND LAGRANGE'S THEOREM	84
3.1	Introduction	84
3.2	Normal Subgroups and Quotient Groups	96

UNIT NO.	SCHEME OF LESSONS	PAGE NO.
Unit – 4	HOMOMORPHISMS	105
4.1	Introduction	105
4.2	Types of Homomorphisms	105
4.3	Isomorphism	112
Unit – 5	CAYLEY'S THEOREM	123
5.1	Cayley's Theorem	123
5.2	Fundamental Theorem of Homomorphism	131
5.3	Rings	132
5.4	Elementary Properties of Rings	138
Unit – 6	VECTOR SPACES	141
6.0	Introduction	141
6.1	Definition and Examples	142
6.2	Subspaces	151
6.3	Linear Transformation	159
6.4	Fundamental Theorem of Homomorphism	162
Unit – 7	SPAN OF A SET	168
7.1	Span of a Set	168
7.2	Linear Independence	171

UNIT NO.	SCHEME OF LESSONS	PAGE NO.
7.3	Basis and Dimension	178
7.4	Rank and Nullity	193
7.5	Matrix of a Linear Transformation	196
Unit – 8	INNER PRODUCT SPACES	202
8.0	Introduction	202
8.1	Definition and Examples	202
8.2	Orthogonality	210
8.3	Orthogonal complement	217
Unit – 9	MATRICES	221
9.0	Introduction	221
9.1	Algebra of Matrices	221
9.2	Elementary Transformations	259
9.3	Rank of a Matrix	269
9.4	Solving Linear Equations	275
9.5	Cayley's Hamilton Theorem	282
9.6	Eigen Values and Eigen Vectors	292

UNIT NO.	SCHEME OF LESSONS	PAGE NO.
Unit – 10	BILINEAR FORMS	314
10.0	Introduction	314
10.1	Bilinear Forms	314
10.2	Matrix of a Bilinear Form	316
10.3	Quadratic Forms	319
10.4	Reduction to Quadratic Forms	324

UNIT – 1

RELATIONS

1.1 Relations

Introduction

A relation provides a criterion for distinguishing certain ordered pairs from others. For example, consider the relation “less than” in the set N of natural numbers. The ordered pair $(2,3)$ is such that 2 is in the given relation to 3. ie, $2 < 3$, whereas the ordered pair $(3,2)$ is not in the given relation. Thus we can collect all ordered pairs in $N \times N$ which are in the given relation. Clearly, this subset ρ of $N \times N$ is given by

$\rho = \{(x, y) / x, y \in N \text{ and } x < y\}$ and this subset of $N \times N$ specifies the relation completely in the sense that $(x, y) \in \rho$ iff $x < y$.

Definition: 1.1.1

Let A and B be non – empty sets. A subset ρ of $A \times B$ is called a **relation** or a **binary relation** from A to B . If an ordered pair $(a, b) \in \rho$ we say that a is related to b in the given relation and we write $a \rho b$. If $(a, b) \notin \rho$ we say that a is not related to b in the given relation.

Though we have defined a relation strictly in set theoretical terms, a relation on a set can be specified by a verbal phrase or a mathematical symbol in such a way that given an ordered pair (a, b) it must be possible for us to decide whether a is in the given relation to b or not. In what follows we confine ourselves to relations defined from a set S to itself.

Examples:

1. $S = \{1, 2, 3, 4\}$. $\rho = \{(1, 2), (2, 3), (3, 3), (1, 4)\}$ is a relation on S .
2. $S = Z$. $a \rho b$ means $a < b$.
3. $S = Z$. $a \rho b$ means $a \leq b$.
4. $S = Z$. $a \rho b$ means a divides b .
5. $S = Z$. $a \rho b$ means $a \equiv b \pmod{m}$ ie. $a - b$ is a multiple of m .
6. $S = Z$. $a \rho b$ means ab is even.
7. $S = Z$. $a \rho b$ means ab is odd.

Space for Hints

8. $S = \mathbb{Z} \cdot a \rho b$ means ab is a perfect square.
9. $S = \wp(A) \cdot B \rho C$ means $B \subseteq C$.
10. $S = \mathbb{R} \cdot a \rho b$ means $a = b$.
11. $S = \mathbb{R} \cdot a \rho b$ means $a - b$ is an integer.
12. $S = \mathbb{R} \cdot a \rho b$ means $a^3 = b$.
13. $S =$ The set of all lines in the Euclidean plane $\mathbb{R} \times \mathbb{R}$. $a \rho b$ means a is parallel to b .
14. S as in example 13. $a \rho b$ means a is perpendicular to b .
15. $S =$ The set of all students in The Madura College. $a \rho b$ means a knows the name of b .
16. S as in example 15. $a \rho b$ means height of $a =$ height of b .
17. $S =$ The set of all students who appeared for the B.Sc. Mathematics examination of the Madurai Kamaraj University in April, 2012. $a \rho b$ means a and b obtain equal marks in the Modern Algebra paper.
18. $S = \mathbb{R} \cdot a \rho b$ means $|a| = |b|$.
19. $S = \mathbb{C} \cdot a \rho b$ means $|a| = |b|$.
20. $S =$ set of all triangles in a plane. $a \rho b$ means a and b are similar.
21. $S = \{1/2, 1/3, 1/4, 2, 3, 4\}$. $a \rho b$ means $a = 1/b$.
22. $S =$ The set of all points in a plane. $a \rho b$ means distance between a and b is less than 1.
23. Let S be any non – empty set. $x \rho y$ means $x = y$.
24. Let S be any non – empty set. $x \rho y$ for all $x, y \in S$.
25. $S =$ The set of all circles in a plane. $a \rho b$ means a and b have the same radius.
26. $S = \mathbb{N} \cdot a \rho b$ means $a^2 - Aan + 3b^2 = 0$.
27. $S = \mathbb{R} \cdot a \rho b$ means $1 + ab > 0$.

1.2 Types of Relations

Definition: 1.2.1

A relation ρ defined on a set S is said to be symmetric if $a \rho b \Rightarrow b \rho a$.

Examples:

1. Consider example 14 of section 1.1. Let $a \rho b$. Then a is perpendicular to b and hence b is perpendicular to a . Hence $b \rho a$. Thus $a \rho b \Rightarrow b \rho a$. Hence ρ is symmetric.
2. Consider example 7 of section 1.1. Let $a \rho b$. Then ab is odd and hence ba is odd. Therefore $b \rho a$. Thus $a \rho b \Rightarrow b \rho a$. Hence ρ is symmetric.
3. Consider example 12 of section 1.1. Here $2 \rho 8$. But $8 \rho 2$ is not true. Hence ρ is not symmetric.

Definition: 1.2.2

A relation ρ defined on a set S is said to be **reflexive** if $a \rho a$ for all $a \in S$.

Examples:

1. Consider example 13 of section 1.1. Clearly any line a is parallel to itself. Thus $a \rho a$ for all $a \in S$. Hence ρ is reflexive.
2. consider example 7 of section 1.1. Here $a \rho a$ is not true for any even integer. Hence ρ is not reflexive.

Definition: 1.2.3

A relation ρ defined on a set S is said to be **transitive** if $a \rho b$ and $b \rho c \Rightarrow a \rho c$.

Examples:

1. Consider example 3 of section 1.1

Let $a \rho b$ and $b \rho c$.

$\therefore a \leq b$ and $b \leq c$.

$\therefore a \leq c$ and hence $a \rho c$.

$\therefore \rho$ is transitive.

2. Consider example 27 of section 1.1.

Here $2 \rho (-1/4)$ and $(-1/4) \rho (-4)$, but $2 \rho (-4)$ is not true.

Hence ρ is not transitive.

Definition: 1.2.4

A relation ρ defined on a set S is said to be an **equivalence**

relation if it is reflexive, symmetric and transitive.

Example:

1. Consider example 5 of section 1.1. For any $a \in \mathbb{Z}$, $a - a = 0$, which is a multiple of m . Hence $a \rho a$.

$\therefore \rho$ is reflexive.

Let $a \rho b$. Then $a - b = km$ where $k \in \mathbb{Z}$.

$\therefore b - a = (-k)m$ and hence $b \rho a$.

$\therefore \rho$ is symmetric.

Let $a \rho b$ and $b \rho c$.

$\therefore a - b = km$ and $b - c = lm$ where $k, l \in \mathbb{Z}$.

$\therefore a - c = (k + l)m$ and hence $a \rho c$.

$\therefore \rho$ is transitive.

Thus ρ is an equivalence relation.

2. Consider example 13 of section 1.1. For any line, a is parallel to itself and hence $a \rho a$ so that the relation is reflexive. Also if a is parallel to b then b is parallel to a and hence the relation is symmetric. Further if a is parallel to b and b is parallel to c then a is parallel to c so that the relation is transitive.

Hence ρ is an equivalence relation.

3. Consider example 8 of section 1.1. with $S = \mathbb{N}$. $a \in \mathbb{N}$. Then $a.a = a^2$ which is a perfect square. Hence ρ is reflexive.

$\therefore a \rho a$.

Clearly, if ab is a perfect square then ba is also a perfect square.

Hence $a \rho b \Rightarrow b \rho a$. Hence ρ is symmetric.

Now let $a \rho b$ and $b \rho c$.

Then $ab = k^2$ and $bc = \ell^2$ where $k, \ell \in \mathbb{Z}$.

$\therefore ab^2c = k^2\ell^2$

$\therefore ac = (k\ell/b)^2$ and hence $a \rho c$.

Hence ρ is transitive.

Thus ρ is an equivalence relation.

Note:

The reflexive, symmetric and transitive properties of the relations given in section 1.1 are summarised in the following table.

Examples	Reflexive	Symmetric	Transitive
5,8,10,11,13,16,17,18,19,20,23,24,25	Yes	Yes	Yes
22,27	Yes	Yes	No
3,9	Yes	No	Yes
15,26	Yes	No	No
7	No	Yes	Yes
6,14,21	No	Yes	No
2,4	No	No	Yes
1,12	No	No	No

Exercises:

1. Verify which of the relation given in section 2.1 are equivalence relations.
2. In $\mathbb{Z}$ we define $a \rho b$ iff $ab > 0$. Is ρ an equivalence relation?
3. Let A be any non-empty set. In $\wp(A)$ we define $B \rho C$ iff $B \cap C \neq \Phi$. Is ρ an equivalence relation?
4. In $\mathbb{Z}$ we define $a \rho b$ iff $a + b$ is even. Is ρ an equivalence relation?
5. Given an example of a relation which is
 - (a) Reflexive and symmetric but not transitive.
 - (b) Reflexive and transitive but not symmetric
 - (c) Symmetric and transitive but not reflexive.
 - (d) Reflexive but not symmetric or transitive.
 - (e) Symmetric but not reflexive or transitive.
 - (f) Transitive but not reflexive or symmetric.
 - (g) Neither reflexive, nor symmetric nor transitive.
6. Is it possible to define an equivalence relation on any non - empty set?
7. Define two equivalence relation on the set $\{1,2,3\}$.

Answers:

2. Not Reflexive 3. Not reflexive, Not Transitive 4. Yes
6. Yes

Definition: 1.2.5

Let ρ be an equivalence relation defined on a set S . Let $x \in S$. The **equivalence class** $[x]$ determined by the element x is defined by

$$[x] = \{y \in S / x \rho y\}.$$

Since $x \rho x$, $x \in [x]$ so that any equivalence class is non – empty.

Examples:

1. Consider the relation ρ defined on Z by $x \rho y \Leftrightarrow x - y$ is a multiple of 3.

$$\begin{aligned} [0] &= \{y \in Z / y - 0 = 3k \quad \text{where } k \in Z\} \\ &= \{0, \pm 3, \pm 6, \dots, \pm 3k, \dots\} \end{aligned}$$

Similarly

$$[1] = \{3k + 1 / k \in Z\} = \{\dots - 5, -2, 1, 4, 7, \dots\}$$

$$[2] = \{3k + 2 / k \in Z\} = \{\dots - 4, -1, 2, 5, 8, \dots\}$$

$$[3] = \{3k + 3 / k \in Z\} = [0] = [6] = [9] = \dots$$

In fact, it is easy to see that $[0]$, $[1]$, $[2]$ are the only three distinct equivalence classes. Any two distinct equivalence classes are disjoint and the union of all these equivalence classes is equal to Z .

2. Let S denote the set of all books in a library. Suppose the library contains books on Mathematics, Physics, Chemistry and Economics. We define a relation ρ on S by $a \rho b$ iff a and b are books on the same subject. Clearly ρ is an equivalence relation.

The equivalence classes are given by

S_1 = Set of all books on Mathematics.

S_2 = Set of all books on Physics.

S_3 = Set of all books on Chemistry.

S_4 = Set of all books on Economics.

3. Let S denote the set of all subsets of $\{1, 2, 3, 4\}$. In S we define ρ by $A \rho B$ if A and B contain the same number of elements. Clearly ρ is an

equivalence relation. The equivalence classes are given by

$$[\Phi] = [\Phi]$$

$$[\{1\}] = \{\{1\}, \{2\}, \{3\}, \{4\}\}$$

$$[\{1,2\}] = \{\{1,2\}, \{1,3\}, \{1,4\}, \{2,3\}, \{2,4\}, \{3,4\}\}$$

$$[\{1,2,3\}] = \{\{1,2,3\}, \{1,2,4\}, \{1,3,4\}, \{2,3,4\}\}$$

$$[\{1,2,3,4\}] = \{\{1,2,3,4\}\}.$$

Exercises:

1. Define an equivalence relation on Z such that the number of equivalence classes is 10.
2. In the examples given in section 2.1, wherever ρ is an equivalence relation, find the equivalence classes.

Definition: 1.2.6

Let ρ be an equivalence relation defined on a set S . The set of all equivalence classes is called the **quotient set** of S and is denoted by S/ρ .

Definition: 1.2.7

Let S be any set. A collection of pairwise disjoint non-empty subsets of S whose union is S is called a **partition** of S .

Examples:

1. Let $S = \{1,2,3,4,5\}$. The subset $\{1\}, \{2\}, \{3,4\}, \{5\}$ form a partition of S .
2. The set of all singletons of a non – empty set S forms a partition of S .

Theorem: 1.1

Let ρ be an equivalence relation defined on a set S . Then

- i. $a \rho b \Leftrightarrow [a] = [b]$.
- ii. Any two distinct equivalence classes are disjoint.
- iii. S is the union of all the equivalence classes.

ie, The set of all equivalence classes forms a partition of S .

Proof:

(i) Let. $a \rho b$. We shall prove that $[a] = [b]$.

Space for Hints

Let $x \in [a]$. Then $x \rho a$.

Since $a \rho b$, by transitivity, we get $x \rho b$.

Hence $x \in [b]$ so that $[a] \subseteq [b]$.

Similarly $[b] \subseteq [a]$. Hence $[b] = [a]$.

Conversely, let $[a] = [b]$.

Then a and b belong to the same equivalence class. Hence $a \rho b$.

(ii) It is enough if we prove that

$$[a] \cap [b] \neq \Phi \Rightarrow [a] = [b].$$

Let $[a] \cap [b] \neq \Phi$. Let $c \in [a] \cap [b]$.

Then $c \in [a]$ and $c \in [b]$.

$$\therefore c \rho a \text{ and } c \rho b$$

$$\therefore a \rho c \text{ and } c \rho b$$

$$\therefore a \rho b \text{ and hence } [a] = [b].$$

(iii) Since each element a of S is in $[a]$ the union of all equivalence classes is S .

Note:

The above theorem shows that every equivalence relation defined on a set S gives rise to a partition of S . The following theorem deals with the converse situation.

Theorem: 1.2

Any partition of a set S determines an equivalence relation ρ such that the members of the partition are precisely the equivalence classes defined by ρ .

Proof:

If $a, b \in S$, we define $a \rho b \Leftrightarrow a$ and b belong to the same member of the partition. Obviously ρ is reflexive and symmetric. Now let $a \rho b$ and $b \rho c$.

$$a \rho b \Leftrightarrow a \text{ and } b \text{ belong to the same partition set } A.$$

$$b \rho c \Leftrightarrow b \text{ and } c \text{ belong to the same partition set } B.$$

Suppose $A \neq B$.

Since $b \in A$ and $b \in B$, $A \cap B \neq \Phi$. This is a contradiction since any

two partition sets are disjoint. Hence $A = B$.

Thus a and $c \in A$ and so that $a \rho c$.

Hence ρ is transitive. Thus ρ is an equivalence relation.

Now let $a \in S$.

Let A be the unique member of the partition such that $a \in A$.

Then $[a] = A$ (by definition of ρ).

Note:

The equivalence relation ρ defined in the above theorem is called the equivalence relation induced by the given partition.

Problems:

Problem: 1

Find the equivalence relation induced by the partition $\{\{1\}, \{2,3\}, \{4\}\}$ of $S = \{1,2,3,4\}$.

Solution:

The equivalence relation ρ induced by the given partition is given by the following subset of $S \times S$, $\{(1,1), (2,2), (3,3), (2,3), (3,2), (4,4)\}$.

Problem: 2

Find the equivalence relation induced by the partition $\{A, B\}$ of Z where $A = \{0,1,2,\dots\}$, $B = \{-1,-2,-3,\dots\}$.

Solution:

Let $x, y \in Z$.

Then $x \rho y \Leftrightarrow x, y \in A$ or $x, y \in B$.

$\therefore x \rho y \Leftrightarrow x, y \geq 0$ or $x, y < 0$.

Problem: 3

If ρ and σ are equivalence relation defined on a set S , prove that $\rho \cap \sigma$ is an equivalence relation.

Solution:

Let $x \in S$.

Then $x \rho x$ and $x \sigma x$ (since ρ, σ are reflexive).

$\therefore x(\rho \cap \sigma)x$. Hence $\rho \cap \sigma$ is symmetric.

Space for Hints

Let $x(\rho \cap \sigma)y$.

Then $x \rho y$ and $x \sigma y$.

$\therefore y \rho x$ and $y \sigma x$ (since ρ, σ are symmetric).

$\therefore y(\rho \cap \sigma)x$. Hence $\rho \cap \sigma$ is symmetric.

Let $x(\rho \cap \sigma)y$ and $y(\rho \cap \sigma)z$.

Then $(x \rho y$ and $x \sigma y)$ and $(y \rho z$ and $y \sigma z)$.

$\therefore (x \rho y$ and $y \rho z)$ and $(x \sigma y$ and $y \sigma z)$.

$\therefore x \rho z$ and $x \sigma z$ (since ρ, σ are transitive).

$\therefore x(\rho \cap \sigma)z$. Hence $\rho \cap \sigma$ is transitive.

Problem: 4

Show that the union of two equivalence relations need not be an equivalence relation.

Solution:

Let $S = \{1, 2, 3\}$.

Let $\rho = \{(1, 1), (2, 2), (3, 3), (1, 2), (2, 1)\}$

and $\sigma = \{(1, 1), (2, 2), (3, 3), (2, 3), (3, 2)\}$.

Clearly ρ and σ are equivalence relations on S . Now $\rho \cup \sigma = \{(1, 1), (2, 2), (3, 3), (1, 2), (2, 1), (2, 3), (3, 2)\}$. $\rho \cup \sigma$ is not transitive since $(1, 2), (2, 3) \in \rho \cup \sigma$ but $(1, 3) \notin \rho \cup \sigma$.

$\therefore \rho \cup \sigma$ is not an equivalence relation.

Problem: 5

What are the smallest and largest equivalence relations on a set S .

Solution:

Any relation on S is a subset $S \times S$. Consider the subset Δ of $S \times S$ given by $\Delta = \{(x, x) / x \in S\}$. Clearly Δ is an equivalence relation on S . Now let ρ be any other equivalence relation on S . Since ρ is reflexive, ρ contains Δ . Hence Δ is the smallest equivalence relation on S .

Obviously the largest equivalence relation on S is given by the subset $S \times S$.

Problem: 6

Let A be a set with n elements.

- i. Find the number of relations that can be defined on A.
- ii. Find the number of reflexive relations that can be defined on A.

Solution:

- i. Any relation on A is a subset of $A \times A$. Since A has n elements, $A \times A$ has n^2 elements.
 $\therefore$ Number of relations that can be defined on A = number of subsets of $A \times A = 2^{n^2}$
- ii. Let $\Delta = \{(a, a) / a \in A\}$. Any reflexive relation A is of the form $\Delta \cup B$ where B is any subset of $(A \times A) - \Delta$ has $n^2 - n$ elements.
 $\therefore$ Number of reflexive relations on
 $A = \text{number of subsets of } (A \times A) - \Delta = 2^{n^2 - n}$.

Exercises:

1. Let ρ be an equivalence relation defined on S. Prove that each element of S belongs to exactly one equivalence class.
2. Prove that $[a] \cap [b] = \Phi \Leftrightarrow a$ is not related to b .
3. Prove that $[a] \subseteq [b] \Leftrightarrow [a] = [b]$.
4. Find the equivalence relations induced by the following partitions of the set $S = \{1, 2, 3, 4\}$.
 (a) $\{1\}, \{2\}, \{3\}, \{4\}$.
 (b) $\{1, 2\}, \{3, 4\}$.
 (c) $\{1, 2, 3, 4\}$.
5. Let $S = \{1, 2, 3, 4\}$. Let $A = \{1, 2\}, B = \{2, 3\}, C = \{1, 4\}$. Define $a \rho b \Leftrightarrow a$ and b both belong to one of the sets A, B or C. Is ρ an equivalence relation?

Partial order

Another type of relation which often arises in mathematics is partial ordering relation for which $\leq$ in R and $\subseteq$ in a power set are examples.

Definition: 1.2.8

A relation ρ defined on a set S is said to be **antisymmetric** if

$$a \rho b \text{ and } b \rho a \Rightarrow a = b.$$

A relation ρ on S which is reflexive, antisymmetric and transitive is called a partial order on S . A set S with a partial order ρ defined on it is called a **partially ordered set** or a **poset** and is denoted by (S, ρ) .

As we have already pointed out $(\mathbb{R}, \leq)$ and $(\wp(S), \subseteq)$ are posets. In $\mathbb{N}$ define $a \rho b \Leftrightarrow a$ divides b . Then ρ is a partial order on $\mathbb{N}$.

Definition: 1.2.9

Let (S, ρ) be a poset. Let $a, b \in S$. a and b are said to be comparable if either $a \rho b$ or $b \rho a$.

Remarks:

In a poset, there may be pairs of elements which are not comparable. In $(\wp(\mathbb{N}), \subseteq)$, $\{1, 2\}$ and $\{1, 3\}$ are not comparable. However in $(\mathbb{R}, \leq)$ any two elements are comparable.

Definition: 1.2.10

A partial order in which any two elements are comparable is called a **linear order** or **total order**. A set S with a linear order defined in it is called a **linearly ordered set** or a **totally ordered set** or a **chain**.

Functions:

The notion of a function is basic in all branches of mathematics. In elementary calculus y is said to be function of x if when x is given, y is determined uniquely. Thus $y = x^2$ is a simple example of a function and this function is given by a rule in the form of an algebraic expression. This function is defined for all real numbers. However the functions $y = 1/x$ and $y = \log x$ are defined only for some values of the variable x . The following general definition includes all such examples.

Definition: 1.2.11

Let A and B be non-empty sets. A **function** or a **mapping** f from A into B , written as $f: A \rightarrow B$ is a rule which assigns to each element $a \in A$ a unique element $b \in B$. The element b which corresponds in this way to a given element $a \in A$ is called the **image** of a under f and is written as $f(a)$. Also if $f(a) = b$ then a is called a **pre image** of b under f . A is called the **domain** of f and $\{f(a) / a \in A\}$ is called the **range** of f . Two functions

$f, g : A \rightarrow B$ are said to be **equal** if $f(x) = g(x)$ for all $x \in A$

Examples:

1. Consider the function $f : \mathbb{Z} \rightarrow \mathbb{Z}$ given by $f(x) = 2x$. Clearly the domain of f is $\mathbb{Z}$. The range of f is given by $\{f(x) / x \in \mathbb{Z}\} = \{2x / x \in \mathbb{Z}\} = \mathbb{E}$. The pre image of any even integer $2n$ under f is n and any odd integer does not have a pre image under f .
2. Consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = x^2$. Any positive real number x has two pre - images under f given by $\sqrt{x}$ and $-\sqrt{x}$ and any negative real number x does not have a pre - image under f . Hence the range of f is $\mathbb{R}^+ \cup \{0\}$.
3. $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = 3$ is called a **constant function**. The range of f is $\{3\}$.
4. $f : \mathbb{Z} \rightarrow \mathbb{E}$ given by $f(x) = 4x$.
5. $f : \mathbb{N} \rightarrow \mathbb{N}$ given by $f(x) = 2^x$.
6. $f : \mathbb{C} \rightarrow \mathbb{R}$ given by $f(z) = |z|$.
7. Let S be the set of all triangles in the Euclidean plane. Let $f : S \rightarrow \mathbb{R}$ be given by $f(s) = \text{area of the triangle } s$.
8. Let $A = \{1, 2, 3, 4\}$ and $B = \{a, b, c\}$. f_1 and f_2 defined below are functions from A to B .
 - (a) $f_1(1) = a = f_1(2)$, $f_1(3) = c = f_1(4)$.
 - (b) $f_2(1) = a = f_2(3)$; $f_2(2) = b$; $f_2(4) = c$.
9. Let $A = \{1, 2, 3\}$ and $B = \{a, b, c\}$. f defined by $f(1) = a, f(2) = b, f(3) = c$ is a function from A to B .
10. Let $A = \{1, 2, 3\}$ and $B = \{a, b, c, d\}$. f defined by $f(1) = a, f(2) = b, f(3) = c$ is a function from A to B .
11. Let A and B be non - empty sets. $f : A \times B \rightarrow A$ defined by $f((a, b)) = a$, is called the projection onto the first coordinate. Projection onto the second coordinate can be similarly defined.
12. Let A be any non - empty set. Let ρ be an equivalence relation

defined on A . $f: A \rightarrow A/\rho$ defined by $f(x) = [x]$ is called the **canonical map** or **natural map**.

13. Let $E \subset \mathbb{R}$. The function $\chi_E: \mathbb{R} \rightarrow \mathbb{R}$ defined by

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{cases}$$

Is called the **characteristic function** on E .

Definition: 1.2.12

Let $f: A \rightarrow B$ be a function. The graph f is defined to be $\{(a, f(a)) / a \in A\}$.

A function may be specified by its graph, which is a subset of $A \times B$. Thus a function from A to B is a relation such that **each** element of A is related to **exactly one** element of B .

Remark:

A relation from A to B may fail to be a function in any one of the following ways.

- i. An element $a \in A$ may be related to more than one element in B .
- ii. An element $a \in A$ may not be related to any element in B .

Examples:

1. Let $A = \{1, 3, 5, 7, 9\}$ and $B = \{2, 4, 6, 8\}$. Consider the relation from A to B given by the following subset of $A \times B$.

$$\{(1, 2), (1, 4), (3, 6), (5, 8), (7, 4)\}.$$

This is not a function from A to B , since 1 is related to 2 and 4. Further 9 is not related to any element of B .

2. The following relation defined on $\mathbb{R}$ by $\{(x, \cos^{-1} x) / x \in \mathbb{R}\}$ is not a function, since $x = 0$ is related to more than one element.

Definition: 1.2.13

A function $f: A \rightarrow B$ is **one – one (injective)** if distinct elements in A have distinct images in B under f . In other words f is 1 – 1 if $x, y \in A$ and $x \neq y \Rightarrow f(x) \neq f(y)$ or equivalently

$$f(x) = f(y) \Rightarrow x = y.$$

The mapping f is called **onto (surjective)** if the range of f is equal to B . Thus if f is onto, every element of B has a pre – image in A .

If $f : A \rightarrow B$ is both 1 - 1 and onto then f is called a **bijection**. In this case every element of B has exactly one pre - image in A .

Examples:

1. $f : \mathbb{Z} \rightarrow \mathbb{Z}$ given by $f(x) = 2x$ is 1 - 1, but not onto. For, $f(x) = f(y) \Rightarrow 2x = 2y \Rightarrow x = y$. Hence f is 1 - 1. The element $3 \in \mathbb{Z}$ does not have any pre - image. Hence f is not onto.
2. Consider $f : \mathbb{Z} \rightarrow \mathbb{Z}$ given by $f(x) = x + 3$. f is 1 - 1 for, $f(x) = f(y) \Rightarrow x + 3 = y + 3 \Rightarrow x = y$. Also any element y has $x = y - 3$ as its pre - image under f . Hence f is onto. Hence f is a bijection.
3. $f : \mathbb{R} \rightarrow \mathbb{R}$ given by $f(x) = 2x$ is 1 - 1 and onto. Each element $y \in \mathbb{R}$ has $y/2$ as its pre - image.
4. Consider $f : \mathbb{R} \rightarrow \mathbb{R}^+$ given by $f(x) = e^x$. Clearly $f(x) = f(y) \Rightarrow e^x = e^y \Rightarrow x = y$.
 $\therefore f$ is 1 - 1.
 Also any element $y \in \mathbb{R}^+$ has $x = \log y$ as its pre - image under f .
 $\therefore f$ is onto. Hence f is a bijection.
5. The function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) = 1/(1+x^2)$ is not 1 - 1, since the elements a and $-a$ have the same image under f . Further f is not onto. Since every element ≤ 0 does not have a pre - image.

Definition: 1.2.14

Let $f : A \rightarrow B$ be a function.

Let $S \subseteq A$. The restriction of f to S , denote by $f|S$, is a function from S to B defined by $(f|S)(x) = f(x)$ for all $x \in S$.

Example:

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) = 1/(1+|x|)$. $f| \mathbb{R}^+ : \mathbb{R}^+ \rightarrow \mathbb{R}$, is given by $(f| \mathbb{R}^+)(x) = 1/(1+x)$.

Definition: 1.2.15

Let $f : A \rightarrow B$ and $g : B \rightarrow C$ be two functions. We define the

composite of these functions $g \circ f : A \rightarrow C$ by the rule $(g \circ f)(a) = g(f(a))$ for all $a \in A$.

Examples:

1. If $f : \mathbb{R} \rightarrow \mathbb{R}$ is given by $f(x) = x^2$ and $g : \mathbb{R} \rightarrow \mathbb{R}$ is given by $g(x) = \sin x$, then $(f \circ g)(x) = f(g(x)) = f(\sin x) = (\sin x)^2$ and $(g \circ f)(x) = g(f(x)) = g(x^2) = \sin x^2$. Thus in general $g \circ f \neq f \circ g$.

2. If $f : \mathbb{R} \rightarrow \mathbb{Z}$ is given by $f(x) = [x]$ and $g : \mathbb{Z} \rightarrow \mathbb{N} \cup \{0\}$ is given by $g(n) = |n|$ then $g \circ f : \mathbb{R} \rightarrow \mathbb{N} \cup \{0\}$ is given by $(g \circ f)(x) = g(f(x)) = g([x]) = |[x]|$.

Result:

If $f : A \rightarrow B$, $g : B \rightarrow C$ and $h : C \rightarrow D$ are functions then $h \circ (g \circ f) = (h \circ g) \circ f$.

Result:

Let $f : A \rightarrow B$, $g : B \rightarrow C$ be bijections. Then $g \circ f : A \rightarrow C$ is also a bijection.

Result:

Let $f : A \rightarrow B$, $g : B \rightarrow C$ be two functions. Then

- i. $g \circ f$ is 1-1 $\Rightarrow f$ is 1-1.
- ii. $g \circ f$ is onto $\Rightarrow g$ is onto.

Definition: 1.2.16

Let $f : A \rightarrow B$ be a bijection. Then for each $b \in B$, there exists a unique element $a \in A$ such that $f(a) = b$. We now define $f^{-1} : B \rightarrow A$ by $f^{-1}(b) = a$. f^{-1} is called the inverse of the function.

Problems:

Problem: 1

Show that $f : \mathbb{R} \rightarrow \mathbb{R}$ defined by $f(x) = 2x - 3$ is a bijection and find its inverse. Compute $f^{-1} \circ f$ and $f \circ f^{-1}$.

Solution:

$f(x) = f(y) \Rightarrow 2x - 3 = 2y - 3 \Rightarrow x = y$. Hence f is 1-1.

Now, let $y \in \mathbb{R}$. To prove that f is onto, we must find $x \in \mathbb{R}$ such that $f(x) = y$.

$$\text{Now } f(x) = y \Rightarrow 2x - 3 = y \Rightarrow x = (y + 3)/2.$$

Hence $(y + 3)/2$ is the pre-image of y under f .

$\therefore$ f is onto. Hence f is a bijection and $f^{-1} : \mathbb{R} \rightarrow \mathbb{R}$ is given by $f^{-1}(x) = (x + 3)/2$.

Now

$$\begin{aligned}(f^{-1} \circ f)(x) &= f^{-1}(f(x)) = f^{-1}(2x - 3) \\ &= [(2x - 3) + 3]/2 = x.\end{aligned}$$

Similarly $(f \circ f^{-1})(x) = x$.

Problem: 2

Show that $f : \mathbb{R} - \{3\} \rightarrow \mathbb{R} - \{1\}$ given by $f(x) = \frac{x-2}{x-3}$ is a bijection and find its inverse.

Solution:

$$\begin{aligned}f(x) = f(y) &\Rightarrow \frac{x-2}{x-3} = \frac{y-2}{y-3} \\ \Rightarrow xy - 3x - 2y + 6 &= xy - 3y - 2x + 6 \\ \Rightarrow 3x + 2y &= 2x + 3y \\ \Rightarrow x &= y.\end{aligned}$$

Hence f is 1-1.

Now, let $y \in \mathbb{R} - \{1\}$. To prove that f is onto we must find $x \in \mathbb{R} - \{3\}$ such that $f(x) = y$.

$$\begin{aligned}\text{Now } f(x) = y &\Rightarrow \frac{x-2}{x-3} = y \\ \Rightarrow x - 2 &= xy - 3y \\ \Rightarrow x &= \frac{2-3y}{1-y}.\end{aligned}$$

Hence $\frac{2-3y}{1-y}$ is the pre-image of y under f .

$\therefore$ f is onto.

Hence f is a bijection and $f^{-1}: \mathbb{R} - \{1\} \rightarrow \mathbb{R} - \{3\}$ is given by

$$f^{-1}(x) = \frac{2-3x}{1-x}.$$

Problem: 3

Show that $f: \mathbb{R} \rightarrow (0,1)$ defined by

$$f(x) = \frac{1}{2} \left[1 + \frac{x}{1+|x|} \right]$$

is a bijection.

Solution:

Clearly $f(0) = 1/2$.

$$\text{When } x > 0, f(x) = \frac{1}{2} \left[1 + \frac{x}{1+x} \right].$$

Hence $(1/2) < f(x) < 1$.

Similarly when $x < 0, f(x) = \frac{1}{2} \left[1 + \frac{x}{1-x} \right]$ and hence $0 < f(x) < 1/2$.

Hence f maps $(0, \infty)$ to $(1/2, 1)$ and $(-\infty, 0)$ to $(0, 1/2)$.

We first prove that $f: (0, \infty) \rightarrow (1/2, 1)$ is a bijection.

Let $x, y \in (0, \infty)$.

Then $f(x) = f(y)$

$$\Rightarrow \frac{1}{2} \left[1 + \frac{x}{1+x} \right] = \frac{1}{2} \left[1 + \frac{y}{1+y} \right]$$

$$\Rightarrow \frac{x}{1+x} = \frac{y}{1+y}$$

$$\Rightarrow x(1+y) = y(1+x)$$

$$\Rightarrow x = y.$$

Hence f is 1-1.

Now, let $y \in (1/2, 1)$. To prove f is onto we must find $x \in (0, \infty)$ such that $f(x) = y$.

$$\text{Now, } f(x) = y \Rightarrow \frac{1}{2} \left[1 + \frac{x}{1+x} \right] = y$$

$$\Rightarrow \frac{x}{1+x} = 2y - 1$$

$$\Rightarrow x = \frac{2y-1}{2(1-y)}$$

Hence f is onto.

$\therefore f : (0, \infty) \rightarrow (1/2, 1)$ is a bijection.

Similarly $f : (-\infty, 0) \rightarrow (0, 1/2)$ is also a bijection.

Hence $f : \mathbb{R} \rightarrow (0, 1)$ is a bijection.

Definition: 1.2.17

Let A be any set. The function $i_A : A \rightarrow A$ defined by $i_A(x) = x$ for all $x \in A$ is called the **identity function** on A . Thus i_A leaves every element of A fixed.

Definition: 1.2.18

Any function $f : A \rightarrow B$ induces two natural set mappings. If $S \subseteq A$, the image of S under f denoted by $f(S)$ is the subset of B given by $\{f(x) / x \in S\}$. Again if $T \subseteq B$, the inverse image of T under f denoted by $f^{-1}(T)$ is the subset of A given by $\{x / f(x) \in T\}$.

Examples:

1. Let $f : \mathbb{Z} \rightarrow \mathbb{Z}$ be given by $f(x) = 2x$. Then

(a) $f(\{1, 2, 3\}) = \{2, 4, 6\}$.

(b) $f^{-1}(\{1, 3, 5\}) = \Phi$, since there is no element $x \in \mathbb{Z}$ such that $f(x) = 1$ or 3 or 5 .

(c) $f^{-1}(\{2, 3, 5\}) = \{1\}$.

2. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be the constant function given by $f(x) = 3$. Then

$f(S) = \{3\}$ for any nonempty subset S of $\mathbb{R}$ and

$$f^{-1}(T) = \begin{cases} \Phi & \text{if } 3 \notin T \\ \mathbb{R} & \text{if } 3 \in T, \text{ where } T \subseteq \mathbb{R} \end{cases}$$

Note that the image of a non – empty set is non – empty whereas the inverse image of a non – empty set may be empty.

1.3 Binary Operations

A binary operation is a “way of putting two things together”. For example in the set $\mathbb{N}$ of natural numbers we can associate with any two

elements $a, b \in \mathbb{N}$ the unique element $a + b \in \mathbb{N}$. Again with any two sets, $A, B \in \wp(X)$ we can associate the set $A \cup B \in \wp(X)$. Note that $+$ in $\mathbb{N}$ gives rise to the function

$$+ : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N} \text{ given by } (a, b) \rightarrow a + b.$$

Definition: 1.3.1

Let A be non – empty set. A binary operation $*$ on A is a function $*$: $A \times A \rightarrow A$. The image of an ordered pair $(a, b) \in A \times A$ under $*$ is denoted by $a * b$. A set A with a binary operation $*$ defined on it is denoted by $(A, *)$.

Examples:

1. The usual addition, $+$, is a binary operation on $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$.

2. On, $\wp(X), \cup, \cap, \Delta$ and are binary operations.

3. Let $A = \{0, 1, 2\}$. A binary operation $*$ on A is given by $0 * 1 = 1$;

$$0 * 2 = 2 * 0 = 2; 1 * 2 = 2 * 1 = 0; 0 * 0 = 0; 1 * 1 = 2; 2 * 2 = 1.$$

If $*$ is a binary operation on a finite set A containing n elements then the n^2 product $a * b, a, b \in A$ can be conveniently arranged in the form of a table containing n rows and n columns, the product $a * b$ coming in the row along a and in the column along b . Thus the above binary operation on A is given by the table.

$*$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Such a table is known as a **Cayley table**.

4. In any non – empty set A , $*$ defined by $a * b = a$ is a binary operation.

5. Let $M_2(\mathbb{R})$ denote the set of all 2×2 matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where

$a, b, c, d \in \mathbb{R}$. We define addition and multiplication in $M_2(\mathbb{R})$ as follows.

$$\text{Let } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad B = \begin{pmatrix} e & f \\ g & h \end{pmatrix}$$

$$A + B = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix} \quad \text{and}$$

$$AB = \begin{pmatrix} ae+bg & af+bh \\ ce+dg & cf+dh \end{pmatrix}$$

Clearly these are binary operations in $M_2(R)$.

6. In N , $*$ defined by $a * b = a^2b$ is a binary operation.
7. Consider $(N, -)$. We have $2 - 3 \notin N$. This difficulty never arises in $(Z, -)$. We say that N is not closed under subtraction so that subtraction is not a binary operation in N .
8. The set of all odd integers is not closed under addition.

Commutativity:

In $(Z, +)$, $2 + 3 = 5$ and $3 + 2 = 5$. These two equations illustrate the law $a + b = b + a$. However in $(Z, -)$, $2 - 3 = -1$ and $3 - 2 = 1$. Thus we see that the law $a * b = b * a$ may be a valid equation for some binary operations and not for some others. Again consider $(N, *)$ where $a * b = a$. Here $2 * 3 = 2$ and $3 * 2 = 3$.

Definition: 1.3.2

A binary operation $*$ on A is said to be **commutative** or to satisfy the commutative law if $a * b = b * a$ for all $a, b \in A$.

Associativity:

In $(Z, +)$, $(2 + 3) + 4 = 5 + 4 = 9$ and $2 + (3 + 4) = 2 + 7 = 9$. These two equations illustrate the law $(a + b) + c = a + (b + c)$. However in $(Z, -)$, $2 - (3 - 4) = 2 - (-1) = 3$ and $(2 - 3) - 4 = -1 - 4 = -5$. Thus we see that the law $(a * b) * c = a * (b * c)$ may be a valid expression for some binary operations and not for others.

Again consider $(N, *)$ where $a * b = a^2b$.

$$2 * (3 * 4) = 2 * 36 = 144 \quad \text{and} \quad (2 * 3) * 4 = 12 * 4 = 576.$$

For union of sets we have $A \cup (B \cup C) = (A \cup B) \cup C$.

Definition: 1.3.3

A binary operation $*$ on A is said to be **associative** or to satisfy the **associative law** if

Space for Hints

$$(a * b) * c = a * (b * c) \text{ for all } a, b, c \in A.$$

Generalizing the associative law:

The associative law has been stated for three elements at a time. However if “.” is an associative binary operation on A and if $a, b, c, d \in A$ then,

$$[(a, b).c].d = [a.(b.c)].d = a.[(b.c).d] = (a.b).(c.d).$$

[Justify each equality here]. This shows that all possible ways of interpreting the product $a. b. c. d$ are equivalent.

Exercise:

If “.” is an associative binary operation on A and if $a, b, c, d, e \in A$.

Show that

$$a.\{b.[c.(d.e)]\} = \{[(a.b).c].d\}.e$$

The associative law gives the product $a. b. c$ a well defined meaning. We have proved that the associative law gives the product $a. b. c. d$ a well defined meaning. The above exercise is a step towards showing that the product $a. b. c. d. e$ has a well defined meaning. This idea may be extended by an inductive argument to prove that the product $a_1.a_2.....a_n$ of a finite number of elements is well defined without the insertion of brackets. Thus if a binary operation is associative brackets can be left out. However brackets cannot be let out if (i) the operation is not associative, for example $2 - (3 - 4) \neq (2 - 3) - 4$, (ii) two operations are involved; for example $2.(3 + 4) \neq 2.3 + 4$.

Exercises:

1. Find all the binary operations on the set $A = \{1, 2\}$.
2. Which of the following sets are closed under the operations indicated? Decide which of the binary operations are commutative and associate.

- (a) $(\mathbb{N}, \text{usual multiplication})$
- (b) $(\mathbb{N}, \text{division})$
- (c) $(\mathbb{Z}, \text{usual multiplication})$
- (d) $(\mathbb{R}, *)$ where $a*b = \min \{a, b\}$

- (e) $(\mathbb{Q}, +)$
 (f) $(\mathbb{R}, +)$
 (g) $(\mathbb{Q}, -)$
 (h) $(\mathbb{Q}^*, \text{division})$
 (i) $(\mathbb{Q}, \text{usual multiplication})$
 (j) $(\mathbb{N}, *)$ where $a * b = 3$
 (k) $(\mathbb{Q}^+, +)$
 (l) $(\mathbb{Q}^+, -)$
 (m) $(\mathbb{N}, *)$ where $a * b = a^b$
 (n) $(\mathbb{Z}, *)$ where $a * b = a^b$
 (o) $(\mathbb{C}, *)$ where $a * b = |a|^b$
 (p) $A = \{1, 2\}$ and $*$ is given by the Cayley table

$*$	1	2
1	1	2
2	2	4

- (q) $A = \{a, b\}$ and $*$ is given by the Cayley table

$*$	a	b
a	a	b
b	a	b

- (r) $(\wp(A), -)$
 (s) $(A^A, *)$ where $f * g = f \circ g$
 (t) $(\mathbb{R} \times \mathbb{R}, *)$ where $(a, b) * (c, d) = (a + b, c + d)$
 (u) $(\mathbb{Z}, *)$ where $a * b = a + b - ab$
 (v) $(\mathbb{E}, *)$ where $a * b = ab/2$
 (w) $(\mathbb{Z}, *)$ where $a * b = a + b + 1$
 (x) $(\mathbb{Z}, *)$ where $a * b = \text{g.c.d of } a \text{ and } b$
 (y) $(\mathbb{Z}, *)$ where $a * b = \text{l.c.m of } a \text{ and } b$
 (z) $S = \{a + b\sqrt{2}/a, b \in \mathbb{Z}\}$, usual addition.
 3. (a) Prove that in $M_2(\mathbb{R})$, matrix addition is commutative and associative.

(b) Prove that in $M_2(\mathbb{R})$, matrix multiplication is associative but not commutative.

In $\mathbb{N}$, under ordinary multiplication, the number $1 \in \mathbb{N}$ is such that $1a = a1 = a$ for all $a \in \mathbb{N}$. In $(\mathbb{Z}, +)$, $0 \in \mathbb{Z}$ is such that $0 + a = a + 0 = a$, for all $a \in \mathbb{Z}$.

Definition: 1.3.4

Let G be a set with a binary operation $*$ defined on it. An element $e \in G$ is called an **identity element** if $e*a = a*e = a$ for all $a \in G$. (An identity element is sometimes called **unit element**. If the operation is addition it is also called **zero element**).

For example, $\mathbb{Q}$ under usual multiplication has 1 as identity. $(\mathbb{Q}, +)$ has 0 as identity. In $(\mathcal{P}(A), \cap)$, $B \cap A = A \cap B = B$ for all $B \subseteq A$ so that A is an identity. For $(\mathcal{P}(A), \cup)$, Φ is an identity.

Exercises:

1. State the identity elements, if they exist, for the binary operations given in the exercise of section 1.3.
2. Prove that $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ are respectively the identity elements for matrix addition and multiplication in $M_2(\mathbb{R})$.

Definition: 1.3.5

Let $*$ be a binary operation defined on G and let e be an identity element for $(G, *)$. Let $a \in G$. An element $b \in G$ is said to be an **inverse of a** if $a*b = b*a = e$.

It is clear from the definition that if a is an inverse of b then b is an inverse of a .

Exercises:

1. What is the inverse of 2 in (a) $\mathbb{Q}$ under usual multiplication? (b) $\mathbb{Q}$ under addition?
2. Show that in $(\mathbb{Z}, +)$ every element has an inverse.
3. What are the elements in $\mathbb{Z}$ which have inverse under usual multiplication?

4. Does every element in $\mathbb{Q}$ have an inverse under usual multiplication?

5. Prove that $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has an inverse in $M_2(\mathbb{R})$ with respect

to matrix multiplication iff $ad - bc \neq 0$ and

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

1.4 Peano's Postulates

1.4.0 Introduction:

We are familiar with the set $N = \{1, 2, 3, \dots, n, \dots\}$ of natural numbers which we use in the process of counting objects. Intuitively the set N may be described as follows. N contains an initial number 1 and to each number in N there is a "successor". Thus the successor of 1 is 2, the successor of 2 is 3 and so on. In other words there is a successor function $\sigma: N \rightarrow N$ and the set N is "generated" from the initial element 1 and the successor function σ . Keeping this in mind, we shall formally describe the set N by a set of axioms due to Peano. We then proceed to construct the set Z of integers and $\mathbb{Q}$ of rational numbers and deduce systematically the familiar properties of Z and $\mathbb{Q}$. Also we indicate without proof the properties that characterize $\mathbb{R}$.

1.4.1 The natural numbers:

Definition: (Peano's postulates)

The system of natural numbers is a set N satisfying the following axioms.

P_1 : There exists an element $1 \in N$.

P_2 : There exists a function $\sigma: N \rightarrow N$ called **successor** function which is one - one.

P_3 : $1 \notin \sigma(N)$ (ie), 1 is not the successor of any element.

P_4 : If $S \subseteq N$ is such that $1 \in S$ and $n \in S \Rightarrow \sigma(n) \in S$, then $S = N$.

Axiom P_4 is called the **principle of induction**. From these postulates we derive the properties of addition, multiplication and order in N .

Space for Hints

Proof:

Let $S = \{x / a + x = b + x \Rightarrow a = b \forall a, b \in N\}$

$1 \in S$, since $a + 1 = b + 1 \Rightarrow a' = a' \Rightarrow a = b$.

Now, let $x \in S$. Then $a + x = b + x \Rightarrow a = b$

Now, $a + x' = b + x' \Rightarrow (a + x)' = (b + x)'$

$$\Rightarrow a + x = b + x$$

$$\Rightarrow a = b.$$

$\therefore x' \in S$. Hence by P_4 , $S = N$.

Exercise:

Prove the **left cancellation law** w.r.t. addition in N , namely,
 $x + a = x + b \Rightarrow a = b$.

Definition: 1.4.2

Let $a \in N$. We define a binary operation called **multiplication** $(.)$ in N as follows.

$$M_1 : a.1 = a$$

$$M_2 : \text{If } a.b \text{ is defined in } N, \text{ then we define } a.b' = a.b + a.$$

Note: 1

$a.b$ is defined for all pairs $a, b \in N$ and multiplication defined as above is unique.

Note: 2

Hereafter we will write ab for $a.b$.

Definition: 1.4.3

Let A be any non - empty set. Let '+' and '.' be two binary operations defined on A . If $a.(b + c) = a.b + a.c$, for all $a, b, c \in A$ then we say that '.' is **left distributive** over "+".

Similarly we say that '.' is **right distributive** over "+" if $(a + b).c$. For example in $\wp(A)$, $\cup$ and $\cap$ are two binary operations which distribute each other, since

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \text{ and}$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$$

Theorem: 1.3

$$(a + b)c = ac + bc \quad \forall a, b, c \in N.$$

Proof:

Let $S = \{c / (a + b)c = ac + bc \ \forall \ a, b \in N\}$

By $M_1, (a + b)1 = a1 + b1$ and hence $1 \in S$.

Now, let $c \in S$. Then $(a + b)c = ac + bc$

Now, $(a + b)c' = (a + b)c + (a + b)$ (by M_2)

$$= (ac + bc) + (a + b)$$

$$= [(ac + bc) + a] + b$$

$$= [ac + (bc + a)] + b$$

$$= [ac + (a + bc)] + b$$

$$= [(ac + a) + bc] + b$$

$$= [(ac + a) + bc] + b$$

$$= (ac + a) + (bc + b)$$

$$= ac' + bc' \quad (\text{by } M_2)$$

$\therefore c' \in S$. Hence by $P_4, S = N$.

Exercises:

1. Prove that multiplication in N is commutative and associative.
2. Prove that $a(b + c) = ab + ac$ for all $a, b, c \in N$.

Definition: 1.4.4

Let $a, b \in N$. We say that a is **greater than** b and write $a > b$ if there exists $p \in N$ such that $a = b + p$. If $a > b$ we also say that b is less than a and we write $b < a$.

Theorem: 1.4

$a < b, b < c \Rightarrow a < c \ \forall \ a, b, c \in N$. (ie) ' $<$ ' is transitive.

Proof:

$a < b \Rightarrow$ there exists $p \in N$ such that $a + p = b$ and

$b < c \Rightarrow$ there exists $q \in N$ such that $b + q = c$

$$\therefore c = b + q = (a + p) + q = a + (p + q).$$

Hence $a < c$.

Theorem: 1.5**Law of trichotomy**

If $a, b \in N$ then exactly one of $a = b, a < b, b < a$ holds.

Proof:

Let $S = \{b / a = b \text{ or } a < b \text{ or } b < a \forall a \in N\}$

Let $a \in N$. Then either $a = 1$ or $a = 1 + k$ for some $k \in N$. Hence either $a = 1$ or $a > 1$. Thus $1 \in S$.

Note:

If $n \in N$ we denote its successor element $\sigma(n)$ by n' .

Theorem: 1.6

The element 1 is the only non – successor element.

Proof:

Let N_1 be the range of the successor function σ . By P_3 , $1 \notin N_1$.

Consider $S = N_1 \cup \{1\}$. Obviously $1 \in S$.

Now, let $y \in S$. Then $y' = \sigma(y) \in \sigma(N) = N_1$.

$\therefore y' \in S$. Hence by P_4 , $S = N$.

$\therefore 1$ is the only non – successor element.

Remark:

Hereafter we denote $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 4$ etc. Thus we get the set $N = \{1, 2, 3, 4, \dots\}$ generated by the element 1 and the function σ .

Definition: 1.5.1

Let $a, b \in N$. We define a binary operation called **addition** (+) in N as follows.

$$A_1 : a + 1 = a'$$

$$A_2 : \text{If } a + b \text{ is defined in } N \text{ then we define } a + b' = (a + b)'$$

Note:

Let $S = \{b / a + b \text{ is defined for all } a \in N\}$ By A_1 , $1 \in S$. By A_2 , $b \in S \Rightarrow b' \in S$. Hence by P_4 , $S = N$. Thus $a + b$ is defined for all pairs $a, b \in N$.

$$\text{For example, } 4 + 2 = 4 + 1' = (4 + 1)' = (4') = 5' = 6.$$

Theorem: 1.7

Addition defined above is **unique**.

Proof:

If possible let $\oplus$ be another operation in N satisfying A_1 and A_2 .

Let $S = \{b / a \oplus b = a + b \text{ for all } a \in N\}$.

By $A_1, a \oplus 1 = a = a + 1$ so that $1 \in S$.

Now, $b \in S \Rightarrow a \oplus b = a + b$

$$\Rightarrow (a \oplus b)' = (a + b)'$$

$$\Rightarrow a \oplus b' = a + b'$$

$$\Rightarrow b' \in S$$

Hence by $P_4, S = N$ so that $a + b = a \oplus b$ for all $a, b \in N$.

Theorem: 1.8

Addition in N is **associative**.

Proof:

Let $S = \{c / a + (b + c) = (a + b) + c, \forall a, b \in N\}$

Now, $a + (b + 1) = a + b' = (a + b)' = (a + b) + 1$. Hence $1 \in S$.

Now, let $c \in S$, so that $a + (b + c) = (a + b) + c$.

Then $a + (b + c') = a + (b + c)'$

$$= [a + (b + c)]'$$

$$= [(a + b) + c]'$$

$$= (a + b) + c'$$

$\therefore c' \in S$. Hence by $P_4, S = N$.

Thus $a + (b + c) = (a + b) + c \forall a, b, c \in N$.

Theorem: 1.9

Addition in N is **commutative**.

Proof:

We shall first prove that $a + 1 = 1 + a \forall a \in N$.

Let $S = \{a / a + 1 = 1 + a\}$

Trivially $1 \in S$. Let $a \in S$. Then $1 + a = a + 1$.

Now, $a' + 1 = (a + 1) + 1 = (1 + a) + 1 = (1 + a)' = 1 + a'$.

$\therefore a' \in S$. Hence by $P_4, S = N$.

$\therefore 1 + a = a + 1$ for all $a \in N$ (1)

Now, let $S_1 = \{b / a + b = b + a \forall a \in N\}$.

By (1), $1 \in S_1$. Let $b \in S_1$ so that $a + b = b + a$.

Then $a + b' = (a + b)'$

$$\begin{aligned}
 &= (b + a)' \\
 &= (b + a) + 1 \\
 &= b + (a + 1) && \text{(by theorem 3)} \\
 &= b + (1 + a) && \text{(by (1))} \\
 &= (b + 1) + a && \text{(by theorem 3)}
 \end{aligned}$$

$\therefore b' \in S_1$. Hence by $P_4, S_1 = N$.

Thus $a + b = b + a$ for all $a, b \in N$.

Theorem: 1.10

$a + x = b + x \Rightarrow a = b$. (This known as **right cancellation law** w.r.t. addition in N).

Now, suppose that $b \in S$.

Then $b = a$, or $b < a$ or $b > a$ for all $a \in N$.

If $b = a$, then $b' = a + 1 > a$.

If $b > a$, then $b = a + p$ for some $p \in N$.

Hence $b' = a + p'$ and thus $b' > a$.

If $b < a$ then $a = b + q$ for some $q \in N$.

If $q = 1, a = b'$, if $q \neq 1$, then $q = r'$ for some $r \in N$, so that $a = b + r' = (b + r)' = (r + b)' = r + b'$.

Hence $b' < a$. In all cases $b' \in S$. Hence by $P_4, S = N$.

Thus if $a, b \in N$ atleast one of the $a = b, a < b, a > b$ holds.

Now we shall prove that exactly one of these statements holds.

If $a = b$ and $a < b$ then $a = b = a + p$ for some $p \in N$. Hence $a = a + p$.

Similarly if $a = b$ and $a > b$ then $a = b = b + q$ for some $q \in N$. If $a < b$ and $a > b$ then $a + r = b$ and $b + s = a$ for some r and $s \in N$, so that $a = b + s = (a + r) + s = a + (r + s)$. Thus all these three cases lead to a contradiction to the fact that $a \neq a + b$ for all $a, b \in N$. Hence the theorem.

Theorem: 1.11

$$a > b \Rightarrow a + c > b + c \quad \forall c \in N.$$

$$a > b \Rightarrow a = b + p \text{ for some } p \in N.$$

$$\therefore a + c = (b + p) + c = b + (p + c) = b + (c + p) = (b + c) + p$$

Hence $a + c > b + c$.

Definition: 1.5.2

Let $a, b \in \mathbb{N}$. We say that a is **less than or equal to** b if $a < b$ or $a = b$ and we write $a \leq b$. $a \leq b$ is also written as $b \geq a$.

Exercises:

Prove the following for natural numbers.

1. $a > b, b > c \Rightarrow a > c$
2. $a \geq b, b \geq c \Rightarrow a \geq c$
3. $a \leq b, b \leq a \Rightarrow a = b$
4. $a \leq b, b < c \Rightarrow a < c$
5. $a \geq b, b > c \Rightarrow a > c$
6. $a < b, c < d \Rightarrow a + c < b + d$
7. $a > b \Leftrightarrow a + c > b + c$
8. $a \geq b \Leftrightarrow a + c \geq b + c$
9. $a \geq 1 \forall a \in \mathbb{N}$
10. $b > a \Rightarrow b \geq a + 1$
11. $a > b \Leftrightarrow ac > bc$
12. $a > b, c > d \Rightarrow ac > bd$
13. $(\mathbb{N}, \leq)$ is a partially ordered set.

Definition: 1.5.3

A partially ordered set $(P, \leq)$ is called a **well ordered set** if any non - empty subset M of P contains an element ℓ and $\ell \leq m$ for all $m \in M$.

Theorem: 1.12

$(\mathbb{N}, \leq)$ is a well ordered set.

Proof:

Let M be a non - empty subset of $\mathbb{N}$.

Let $S = \{n / n \leq m \forall m \in M\}$. Clearly $1 \in S$. Also for any $m \in M$, $m' \notin S$ and hence $S \neq \mathbb{N}$. Hence there exist an element ℓ such that $\ell \in S$ and $\ell' \notin S$. Since $\ell \in S$, $\ell \leq m$ for all $m \in M$. Also if $\ell \notin M$ then $\ell < m$ for all $m \in M$. Hence $\ell' \leq m$ so that $\ell' \in S$ which is a

contradiction. Therefore $\ell \in M$ and is the required least element. Hence N is a well ordered set.

1.5 Principal of Induction:

The axiom of induction P_4 has many important consequences. It gives us two important principles (for proving certain types of problems known as the first and second principles of induction.

First Principle of Induction:

Theorem: 1.13

Let $\{P(n) \mid n \in N\}$ be a set of statements one for each natural number n . If (i) $P(1)$ is true and (ii) $P(k)$ is true $\Rightarrow P(k+1)$ is true, then $P(n)$ is true for all natural numbers n .

Proof:

Let $S = \{s \mid s \in N, P(s) \text{ is true}\}$.

By (i) $P(1)$ is true and hence $1 \in S$. $s \in S \Rightarrow P(s)$ is true $\Rightarrow P(s+1)$ is true (by (ii)) $\Rightarrow s+1 \in S \Rightarrow s' \in S$.

$\therefore$ by the axiom of induction $P_4, S = N$, i.e., $P(n)$ is true for all natural numbers n .

Many theorems are proved more easily by induction method than by direct methods. The proof consists of two stages, namely, (i) establishing that the statement $P(n)$ under consideration is true for $n = 1$; i.e., $P(1)$ is true, (ii) assuming that $P(k)$ is true (k is any natural number) and showing that $P(k+1)$ is true.

Problems:

1. Prove by induction:

$$(a) \quad 1 + 2 + \dots + n = \frac{n(n+1)}{2}.$$

$$(b) \quad 1 + x + x^2 + \dots + x^{n-1} = \frac{x^n - 1}{x - 1} \text{ when } x > 1.$$

$$(c) \quad 11 + 23 + 59 + \dots + (2 \cdot 3^n + 5) = 3^{n+1} + 5n - 3.$$

$$(d) \quad 3 \cdot 5^{2n+1} + 2^{3n+1} \text{ is divisible by } 17.$$

Solution:

(a) Let $P(n) : 1 + 2 + \dots + n = \frac{n(n+1)}{2}$.

Stage: 1

When $n = 1$, L.H.S = 1 and R.H.S = $\frac{1(1+1)}{2} = 1$

$\therefore P(1)$ is true.

Stage: 2

Assume that $P(k)$ is true

i.e., $1 + 2 + \dots + k = \frac{k(k+1)}{2}$.

(This is known as induction hypothesis).

To prove that $P(k+1)$ is true

i.e., $1 + 2 + \dots + (k+1) = \frac{(k+1)(k+1+1)}{2}$

Now $1 + 2 + \dots + (k+1)$

$= [1 + 2 + \dots + k] + (k+1)$

$= \frac{k(k+1)}{2} + (k+1)$ by induction hypothesis.

$= \frac{(k+1)}{2} (k+2)$

$= \frac{(k+1)(k+1+1)}{2}$

$\therefore P(k+1)$ is true.

Thus by (first principle of) induction $P(n)$ is true for all natural numbers n .

i.e., $1 + 2 + \dots + n = \frac{n(n+1)}{2}$

(b) Let $P(n) : 1 + x + x^2 + \dots + x^{n-1} = \frac{x^n - 1}{x - 1}$

When $x > 1$.

When $n = 1$, L.H.S = 1 and R.H.S = $\frac{x - 1}{x - 1} = 1$.

$\therefore P(1)$ is true.

Assume that $P(k)$ is true

$$\text{i.e., } 1+x+\dots+x^{k-1} = \frac{x^k - 1}{x - 1} \quad (1)$$

$$\begin{aligned} & 1+x+x^2+\dots+x^{k+1-1} \\ &= (1+x+\dots+x^{k-1})+x^k \\ &= \frac{x^k - 1}{x - 1} + x^k \text{ by (1)} \\ &= \frac{x^k - 1 + x^k(x - 1)}{x - 1} \\ &= \frac{x^{k+1} - 1}{x - 1} \end{aligned}$$

This shows that $P(k + 1)$ is true. Therefore, by induction principle it follows

that $P(n)$ is true for all natural numbers n , i.e., $1+x+\dots+x^{n-1} = \frac{x^n - 1}{x - 1}$ for $x > 1$.

(c) Let

$$P(n): 11+23+59+\dots+(2 \cdot 3^n + 5) = 3^{n+1} + 5n - 3.$$

When $n = 1$, L.H.S = 11 and R.H.S = $3^{1+1} + 5 \cdot 1 - 3 = 11$ so that $P(1)$ is true.

Assume that $P(k)$ is true.

$$\text{i.e., } 11+23+\dots+(2 \cdot 3^k + 5) = 3^{k+1} + 5k - 3 \dots \quad (1)$$

$$\begin{aligned} & \text{Then } 11+23+\dots+(2 \cdot 3^{k+1} + 5) \\ &= [11+23+\dots+(2 \cdot 3^k + 5)] + (2 \cdot 3^{k+1} + 5) \\ &= 3^{k+1} + 5k - 3 + (2 \cdot 3^{k+1} + 5) \text{ from (1)} \\ &= 3 \cdot 3^{k+1} + 5(k+1) - 3 \\ &= 3^{n+1} + 5n - 3 \text{ where } n = k+1. \end{aligned}$$

Hence $P(k + 1)$ is true. Thus by induction principle, $P(n)$ is true for all natural numbers n .

(d) Let $P(n): 3 \cdot 5^{3n+1} + 2^{3n+1}$ is divisible by 17. When

$$n = 1, 3 \cdot 5^{3 \cdot 1 + 1} + 2^{3 \cdot 1 + 1} = 3 \cdot 5^4 + 2^4 = 391 = 17 \times 23 \text{ which is divisible by 17.}$$

$\therefore P(1)$ is true. Assume that $P(k)$ is true, i.e., $3 \cdot 5^{2k+1} + 2^{3k+1}$ is divisible by 17.

When $n = k+1$, $3 \cdot 5^{2n+1} + 2^{3n+1}$

$$= 3 \cdot 5^{2k+3} + 2^{3k+4}$$

$$= 3 \cdot 5^{2k+1} \cdot 5^2 + 2^{3k+1} \cdot 2^3$$

$$= 3 \cdot 5^{2k+1} \cdot (17+8) + 2^{3k+1} \cdot 8$$

$$= 3 \cdot 5^{2k+1} \cdot 17 + 8(3 \cdot 5^{2k+1} + 2^{3k+1}) \text{ which is divisible by 17}$$

on using the induction hypothesis (1). Hence $P(k+1)$ is true. Therefore by the principle of induction $P(n)$ is true for all natural numbers n .

2. Prove that

$$\frac{1}{1 \cdot 4} + \frac{1}{4 \cdot 7} + \frac{1}{7 \cdot 10} + \dots + \frac{1}{(3n-2)(3n+1)} = \frac{n}{3n+1}.$$

Proof:

We prove by induction.

Let

$$P(n): \frac{1}{1 \cdot 4} + \frac{1}{4 \cdot 7} + \frac{1}{7 \cdot 10} + \dots + \frac{1}{(3n-2)(3n+1)} = \frac{n}{3n+1}.$$

When $n = 1$, L.H.S. = $\frac{1}{1 \cdot 4} = \frac{1}{4}$ and

$$\text{R.H.S.} = \frac{1}{(3-2)(3+1)} = \frac{1}{4}.$$

$\therefore P(1)$ is true.

Assume that $P(k)$ is true.

$$\text{i.e., } \frac{1}{1 \cdot 4} + \frac{1}{4 \cdot 7} + \dots + \frac{1}{(3k-2)(3k+1)} = \frac{k}{3k+1} \dots (1)$$

$$\text{Then } \frac{1}{1 \cdot 4} + \frac{1}{4 \cdot 7} + \dots + \frac{1}{[3(k+1)-2][3(k+1)+1]}$$

$$= \left[\frac{1}{1 \cdot 4} + \frac{1}{4 \cdot 7} + \dots + \frac{1}{(3k-2)(3k+1)} \right] + \frac{1}{(3k+1)(3k+4)}$$

$$= \frac{k}{3k+1} + \frac{1}{(3k+1)(3k+4)} \text{ from (1)}$$

$$= \frac{k(3k+4)+1}{(3k+1)(3k+4)} = \frac{(3k+1)(k+1)}{(3k+1)(3k+4)}$$

$$= \frac{k+1}{3k+4} = \frac{k+1}{3(k+1)+1}$$

This shows that $P(k+1)$ is true and hence by induction principle it follows that $P(n)$ is true for all natural numbers n .

3. Show by the method of induction that $n^2 > 3(n+1)$ for $n \geq 4$ and hence that $3^n > n^3$ for $n \geq 4$.

Solution:

(i) Say $P(n): n^2 > 3(n+1)$ for $n \geq 4$. When $n=4, n^2=16$ and $3(n+1)=15$. Since $16 > 15$, $P(4)$ is true.

Assume that $P(k)$ is true ($k \geq 4$), i.e., $(k+1)^2 > 3(k+1+1)$.

$$\text{Now } (k+1)^2 = k^2 + 2k + 1$$

$$> 3(k+1) + 2 + 1 \text{ from (1)}$$

$$= 5k + 4$$

$$> 3k + 6 (\because k \geq 4)$$

$$\text{i.e., } (k+1)^2 > 3(k+2).$$

Hence $P(k+1)$ is true. By the principle of induction it follows that $P(n)$ is true for $n \geq 4$.

(ii) Again, let $P(n): 3^n > n^3$ for $n \geq 4$. When $n=4, 3^n=3^4=81$ and $n^3=4^3=64$. Since $81 > 64$, $P(4)$ is true.

Suppose that $P(k)$ is true ($k \geq 4$)

$$\text{i.e., } 3^k > k^3 \quad (2)$$

To show that $3^{k+1} > (k+1)^3$.

$$\text{Now } 3^{k+1} = 3 \cdot 3^k > 3 \cdot k^3 \text{ from (2).}$$

$$\text{But } n^2 > 3(n+1) \text{ for } n \geq 4. \therefore k^2 > 3(k+1) \text{ and hence}$$

$$k^3 > 3k(k+1) = 3k^2 + 3k.$$

$$\therefore k^2 = k^3$$

$$k^3 > 3k^2 + 3k \text{ and}$$

$$k^3 > 1 \quad (\because k \geq 4).$$

$$\text{Adding } 3k^3 > k^2 + 3k^2 + 3k + 1 = (k+1)^3.$$

$\therefore 3^{k+1} > (k+1)^3$, i.e., $P(k+1)$ is true. By the principle of induction it follows that $P(n)$ is true for $n \geq 4$.

4. If S is a set having n elements, show that its power set $\wp(S)$ has 2^n elements.

Solution:

Let $P(n)$ be the statement "if S is a set having n elements, then $\wp(S)$ has 2^n elements".

(i) Let $S = \{x\}$, then $\wp(S) = \{\phi, \{x\}\}$. Hence if S is a set having 1 element, then $\wp(S)$ has 2^1 elements, i.e., $P(1)$ is true.

(ii) Assume that $P(k)$ is true, i.e., if a set has k elements, then it has 2^k subsets. (1)

Let S be a set having $k+1$ elements.

Say $S = \{x_1, x_2, \dots, x_{k+1}\}$. Then the number of subsets of S = number of subsets of S not containing x_{k+1} + number of subsets of S containing x_{k+1} . (2)

Now, number of subsets of S not containing x_{k+1} = number of subsets of the set $T = \{x_1, x_2, \dots, x_k\} = 2^k$ by induction hypothesis (1) since T has k elements.

Also each subset of S containing x_{k+1} can be obtained from some subset of T by adding x_{k+1} to it. Since there are 2^k subsets of T , there are exactly 2^k subsets of S each containing x_{k+1} .

Hence from (2), number of subsets of $S = 2^k + 2^k = 2 \times 2^k = 2^{k+1}$ i.e., $\wp(S)$ has 2^{k+1} elements. $\therefore P(k+1)$ is true.

Thus, by induction it follows that $\wp(n)$ is true for all natural numbers n .

5. Let $P(n)$ be the statement

$$1+3+5+\dots+(2n-1)=n^2+5.$$

Show that $P(k)$ is true $\Rightarrow P(k+1)$ is true. Can we conclude that $P(n)$ is

Space for Hints

true for all n ?

Assume that $P(k)$ is true.

$$\text{That is, } 1+3+\dots+(2k-1)=k^2+5 \quad (1)$$

$$\text{Then } 1+3+5+\dots+[2(k+1)-1]$$

$$=[1+3+5+\dots+(2k-1)]+(2k+1)$$

$$=(k^2+5)+2k+1 \text{ from (1)}$$

$$=(k+1)^2+5.$$

$\therefore P(k+1)$ is true.

What about $P(1)$? When $n=1$, L.H.S=1 and R.H.S= $1^2+5=6$, i.e., $P(1)$ is not true. Hence we cannot conclude that $P(n)$ is true for all n .

Note:

The above example shows the importance of the two stages in the proof by mathematical induction.

Problem: 6

Using mathematical induction show that $3^{2^n}-1$ is a multiple of 8.

Solution:

$$\text{Let } P(n)=3^{2^n}-1$$

$$\therefore P(1)=3^2-1=8 \text{ which is a multiple of 8.}$$

Hence the result is true for $n=1$.

Let us assume that the result is true for $n=k$.

$$\therefore P(k)=3^{2^k}-1=8m \text{ for some } m \in \mathbb{N}. \quad (1)$$

$$\text{Now, } P(k+1)=3^{2^{k+1}}-1$$

$$=3^{2^k}3^{2^k}-1$$

$$=(8m+1)3^2-1 \quad (\text{using 1})$$

$$=72m+9-1$$

$$=72m+8$$

$$=8(9m+1) \text{ which is a multiple of 8.}$$

Hence the result is true for $n=k+1$.

$\therefore$ By induction principle the result is true for all $n \in \mathbb{N}$.

Problem: 7

Using mathematical induction show that $3^{2n} + 24n - 1$ is divisible by 32.

Solution:

$$\text{Let } P(n) = 3^{2n} + 24n - 1$$

$$\therefore P(1) = 3^2 + 24 - 1 = 32 \text{ which is a multiple of 32.}$$

Hence the result is true for $n = 1$.

Let us assume that the result is true for $n = k$.

$$\therefore P(k) = 3^{2k} + 24k - 1 \text{ is divisible by 32.} \quad (1)$$

$$\text{Now, } P(k+1) = 3^{2k+2} + 24(k+1) - 1$$

$$= 9(3^{2k} + 24k - 1) - 8 \times 24k + (24 + 9 - 1)$$

$$= 9P(k) - 192k + 32 \text{ which is divisible by 32 by (1)}$$

Hence the result is true for $n = k+1$.

$\therefore$ By induction principle the result is true for all $n \in \mathbb{N}$.

Exercises:

By the principle of mathematical induction prove the following, where $n \in \mathbb{N}$.

$$1. \quad n^2 + n + 1 \text{ is an odd number.}$$

$$2. \quad 1^2 + 2^2 + 3^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$$

$$3. \quad 1^3 + 2^3 + 3^3 + \dots + n^3 = \frac{n^2(n+1)^2}{4}$$

$$4. \quad (x+a)^n = x^n + {}^nC_1 x^{n-1}a + \dots + {}^nC_r x^{n-r}a^r + \dots + a^n$$

$$5. \quad 1^2 + 3^2 + 5^2 + \dots + (2n-1)^2 = \frac{n(4n^2-1)}{3}$$

$$6. \quad 1.2 + 2.3 + 3.4 + \dots + n.(n+1) = \frac{n(n+1)(n+2)}{3}$$

$$7. \quad n^7 - n \text{ is divisible by 42.}$$

$$8. \quad 5^{2n} - 1 \text{ is a multiple of 6.}$$

$$9. \quad 7^{2n} - 1 \text{ is a multiple of 48.}$$

10. $8^{n+1} - 7n + 41$ is a multiple of 49.

11. 64 divides $3^{2n} - 8n - 1$.

12. 360 divides $19^{2n} - 1$.

Second Principle of Induction:

Theorem: 1.14

Let $\{P(n) \mid n \in \mathbb{N}\}$ be a set of statements one for each natural number n . If (i) $P(1)$ is true and (ii) if for each natural number k , $P(m)$ is true for all $m < k$ implies that $P(k)$ is true, then $P(n)$ is true for all natural numbers n .

Proof:

Let $S = \{s \in \mathbb{N} \mid P(s) \text{ is false}\}$.

Since $P(1)$ is true, $1 \notin S$.

We have to show that $S = \emptyset$.

Suppose $S \neq \emptyset$. Then S is a non-empty subset of $\mathbb{N}$ and hence by the well-ordering principle S has a least element ℓ , i.e., $\ell \in S$ and $\ell < s \forall s \in S$.

Observe that $\ell \neq 1$ ($\because 1 \notin S$).

If $m < \ell$, then $m \notin S$ ($\because \ell$ is the least element of S). Hence $P(m)$ is true for each $m < \ell$. Thus by hypothesis (ii) it follows that $P(\ell)$ is true. This means that $\ell \notin S$ which is a contradiction. $\therefore S = \emptyset$. Thus $P(s)$ is true for all natural numbers s .

Remark: 3.1.32

The hypothesis (ii) of the second principle of induction **appears** stronger than the corresponding hypothesis of the first principle of induction. Hence the second principle of induction is sometimes called the **strong** principle of induction while the first principle of induction is called the **weak** principle of induction. However both the principles are equivalent.

UNIT - 2

SUBGROUPS

Introduction: 2.0

Groups: 2.0.1

Modern Algebra is largely concerned with the study of abstract sets endowed with one or more binary operations. In this chapter we introduce one of the basic algebraic structures known as **groups**. A group is a set with one binary operation defined on it satisfying some natural conditions. The definition of a group is an abstraction of the familiar properties of $(\mathbb{Z}, +)$ given below.

- (i) Addition is an associative binary operation in $\mathbb{Z}$.
- (ii) The element $0 \in \mathbb{Z}$ is such that $a + 0 = 0 + a = a$ for all $a \in \mathbb{Z}$. Hence 0 is the identity element w.r.t. addition.
- (iii) Let $a \in \mathbb{Z}$. The element $-a \in \mathbb{Z}$ is such that $a + (-a) = (-a) + a = 0$. Hence $-a$ is the inverse of a .

We isolate these properties in the following definition.

Definition and Examples:

Definition: 2.0.2

A non - empty set G together with a binary operation $*: G \times G \rightarrow G$ is called a **group** if the following conditions are satisfied.

- (i) $*$ is **associative** (i.e.) $a * (b * c) = (a * b) * c$ for all $a, b, c \in G$.
- (ii) There exists an element $e \in G$ such that $a * e = e * a = a$ for all $a \in G$. e is called the **identity** element of G .
- (iii) For any element a in G there exists an element $a' \in G$ such that $a * a' = a' * a = e$. a' is called the **inverse** of a .

Examples:

1. $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are groups under usual addition.

2. The set of all 2×2 matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where $a, b, c, d \in \mathbb{R}$ is a group under matrix addition.

$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ is the identity element and $\begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$ is the inverse of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$

3. The set of all 2×2 non-singular matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ where $a, b, c, d \in \mathbb{R}$ is a group under matrix multiplication.

We know that matrix multiplication is associative.

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ is the identity element.

The inverse of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is $\frac{1}{|A|} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ where $|A| = ad - bc \neq 0$.

4. $\mathbb{N}$ is not a group under usual addition since there is no element $e \in \mathbb{N}$ such that $x + e = x$.

5. The set $\mathbb{E}$ of all even integers under usual addition is a group. For $a, b \in \mathbb{E} \Rightarrow a + b \in \mathbb{E}$. Therefore usual addition is a binary operation in $\mathbb{E}$.

$0 \in \mathbb{E}$ is the identity element.

If $a \in \mathbb{E}$, $-a \in \mathbb{E}$ is the inverse of a .

6. $\mathbb{Q}^*$ and $\mathbb{R}^*$ under usual multiplication are groups. 1 is the identity element and the inverse of a is $1/a$.

7. $\mathbb{Q}^+$ is a group under usual multiplication. For $a, b \in \mathbb{Q}^+ \Rightarrow ab \in \mathbb{Q}^+$. Therefore usual multiplication is a binary operation in $\mathbb{Q}^+$.

$1 \in \mathbb{Q}^+$ is the identity element.

If $a \in \mathbb{Q}^+$, $(1/a) \in \mathbb{Q}^+$ is the inverse of a .

8. $\mathbb{Z}$ under the usual multiplication is not a group. $1 \in \mathbb{Z}$ is the identity element. However any element other than 1 and -1 does not have an inverse.

9. Let $G = \{1, -1\}$. G is a group under usual multiplication. 1 is the identity element. The inverse of each element is itself. The Cayley table for this group is

*	1	-1
1	1	-1
-1	-1	1

10. Let $G = \{1, i, -1, -i\}$. G is a group under usual multiplication. The identity element is 1. The inverse of 1, i , -1 and $-i$ are 1, $-i$, -1 and i respectively.

The Cayley table for this group is given by

*	1	i	-1	-i
1	1	i	-1	-i
i	i	-1	-i	1
-1	-1	-i	1	i
-i	-i	1	i	-1

11. Let $G = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$

G is a group under matrix multiplication. (Construct the Cayley table for this group).

Since $a^2 + b^2 \neq 0, 1/x \in C^*$ and is the inverse of x . Hence C^* is a group under usual multiplication.

12. Let $G = \{z/z \in \mathbb{C} \text{ and } |z|=1\}$. G is a group under usual multiplication.

Proof:

Let $z_1, z_2 \in G$.

Then $|z_1| = |z_2| = 1$.

$\therefore |z_1 z_2| = |z_1| |z_2| = 1$ and hence $z_1 z_2 \in G$.

We know that usual multiplication of complex numbers is associative.

Also $1 = 1 + i0 \in G$ and is the identity element.

Now, let $z \in G$. Then $|z|=1$.

Hence $|1/z| = 1/|z| = 1$.

$\therefore 1/z \in G$ and is the inverse of z .

Hence G is a group.

13. The set of all n^{th} roots of unity with usual multiplication is a group.

Proof:

Let $w = \cos(2\pi/n) + i \sin(2\pi/n)$. Then the n^{th} roots of unity are given by $1, \omega, \omega^2, \dots, \omega^{n-1}$.

Let $G = \{1, \omega, \omega^2, \dots, \omega^{n-1}\}$.

We know that $\omega^n = 1, \omega^{n+1} = \omega$ etc.

Let $\omega^r, \omega^s \in G$. Let $r+s = qn+t$ where $0 \leq t < n$.

$$\therefore \omega^r \omega^s = \omega^{r+s} = \omega^{qn+t} = (\omega^n)^q \omega^t = \omega^t \in G.$$

We know that usual multiplication of complex number is associative.

$1 \in G$ is the identity element.

Inverse of ω^r is ω^{n-r} . Hence G is a group.

14. Let $G = \{a + b\sqrt{2} / a, b \in \mathbb{Z}\}$. Then G is a group under usual addition.

Proof:

Let $a + b\sqrt{2}$ and $c + d\sqrt{2} \in G$. Then

$$(a + b\sqrt{2}) + (c + d\sqrt{2}) = (a + c) + (b + d)\sqrt{2} \in G.$$

We know that usual addition is associative. $0 = 0 + 0\sqrt{2} \in G$ is the identity element.

$-a - b\sqrt{2}$ is the inverse of $a + b\sqrt{2}$. Hence G is group.

15. Let G be the set of all real numbers except -1 . Define $*$ on G by $a * b = a + b + ab$. Then $(G, *)$ is a group.

Proof:

Let $a, b \in G$. Then $a \neq -1$ and $b \neq -1$. We claim that $a + b \neq -1$.

Suppose $a * b = -1$.

Then $a + b + ab = -1$ so that $a + b + ab + 1 = 0$. i.e., $(a+1)(b+1) = 0$ so

that either $a = -1$ or $b = -1$ which is a contradiction. Hence

$a * b \neq -1$ and thus $*$ is a binary operation on G .

$*$ is associative, for $a * (b * c)$

$$= a * (b + c + bc)$$

$$= a + (b + c + bc) + a(b + c + bc)$$

$$= a + b + c + bc + ab + ac + abc.$$

$$\text{Also } (a * b) * c = (a + b + ab) * c$$

$$\begin{aligned}
 &= a + b + ab + c + (a + b + bc)c \\
 &= a + b + c + ab + ac + bc + abc.
 \end{aligned}$$

Hence $a * (b * c) = (a * b) * c$.

0 is the identity, for $a * 0 = a + 0 + 0a = a$ and $0 * a = 0 + a + 0a = a$.

Now, let a' be such that $a * a' = 0$. Hence $a + a' + aa' = 0$ so that $a' = -a/(1 + a)$.

Since $a \neq -1$, we have $a' \in R - \{-1\}$

$$\begin{aligned}
 \text{Also } a' * a &= \frac{-a}{1+a} * a \\
 &= \frac{-a}{1+a} + a + \frac{-a^2}{1+a} = 0.
 \end{aligned}$$

Hence a' is the inverse of a .

Thus G is a group.

16. In R^* we define $a * b = (1/2)ab$. Then $(R^*, *)$ is a group.

Proof:

Obviously $*$ is a binary operation in R^* .

Let $a, b, c \in R^*$.

Then $(a * b) * c = [(1/2)ab] * c = (1/4)abc = a * (b * c)$. Hence $*$ is associative.

Let $e \in R^*$ be such that $a * e = a$.

$$\therefore (1/2)ae = a \text{ and hence } e = 2.$$

$$\therefore 2 * a = a * 2 = a. \text{ Hence } 2 \text{ is the identity.}$$

Let $a \in R^*$. Let $b \in R^*$ be such that $a * b = 2$. Then

$$(1/2)ab = 2, \text{ (i.e.) } b = 4/a.$$

$$\therefore a * (4/a) = 1/2(4/a)a = 2 \text{ i.e., } (4/a) \text{ is the inverse of } a.$$

Thus $(R^*, *)$ is a group.

17. Let G denote the set of all matrices of the form $\begin{pmatrix} x & x \\ x & x \end{pmatrix}$ where $x \in R^*$. Then G is group under matrix multiplication.

Proof:

Let $A, B \in G$. Let $A = \begin{pmatrix} x & x \\ x & x \end{pmatrix}$ and $B = \begin{pmatrix} y & y \\ y & y \end{pmatrix}$.

Then $AB = \begin{pmatrix} 2xy & 2xy \\ 2xy & 2xy \end{pmatrix} \in G$.

We know that matrix multiplication is associative.

Let $E = \begin{pmatrix} e & e \\ e & e \end{pmatrix}$ be such that $AE = A$.

$$\therefore \begin{pmatrix} x & x \\ x & x \end{pmatrix} \begin{pmatrix} e & e \\ e & e \end{pmatrix} = \begin{pmatrix} x & x \\ x & x \end{pmatrix}$$

$$\therefore \begin{pmatrix} 2xe & 2xe \\ 2xe & 2xe \end{pmatrix} = \begin{pmatrix} x & x \\ x & x \end{pmatrix}$$

$$\therefore 2xe = x. \text{ Hence } e = 1/2.$$

Hence $E = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$ is the identity element of G .

Let $\begin{pmatrix} y & y \\ y & y \end{pmatrix}$ be the inverse of $\begin{pmatrix} x & x \\ x & x \end{pmatrix}$. Then

$$\begin{pmatrix} x & x \\ x & x \end{pmatrix} \begin{pmatrix} y & y \\ y & y \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

$$\therefore \begin{pmatrix} 2xy & 2xy \\ 2xy & 2xy \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

$$\therefore 2xy = 1/2. \text{ Hence } y = x/4.$$

$$\therefore \text{Inverse of } \begin{pmatrix} x & x \\ x & x \end{pmatrix} \text{ is } \begin{pmatrix} x/4 & x/4 \\ x/4 & x/4 \end{pmatrix}.$$

Hence G is group.

18. Let $G = \{(a, b) / a \in \mathbb{R}^*, b \in \mathbb{R}\}$. Then G is a group under the operation $*$ defined by $(a, b) * (c, d) = (ac, bc + d)$.

Proof:

Clearly $*$ is a binary operation defined on G . Now, let $x = (a, b)$, $y = (c, d)$ and $z = (e, f)$ be three elements in G .

$$\begin{aligned} \text{Then } x * (y * z) &= (a, b) * (ce, de + f) \\ &= (ace, bce + de + f) \end{aligned}$$

$$\text{Also } (x * y) * z = (ac, bc + d) * (e, f)$$

$$= (ace, bce + de + f).$$

$$\therefore x * (y * z) = (x * y) * z, \text{ so that } * \text{ is associative.}$$

Now, to find the identity element in G suppose

$$(a, b) * (c, d) = (a, b).$$

$$\text{Then } (ac, bc + d) = (a, b).$$

$$\therefore ac = a \text{ and } bc + d = b.$$

$$\text{Hence } c = 1 \text{ and } d = 0.$$

$$\text{Thus } (a, b) * (1, 0) = (1, 0) * (a, b) = (a, b).$$

$$\therefore (1, 0) \in G \text{ is identity element.}$$

$$\text{Now suppose } (a, b) * (a', b') = (1, 0).$$

$$\therefore (aa', ba' + b') = (1, 0).$$

$$\text{Hence } aa' = 1 \text{ and } ba' + b' = 0$$

$$\therefore a' = 1/a \text{ and } b' = -b/a.$$

$$\text{Thus } (a, b) * (1/a, -b/a) = (1/a, -b/a) * (a, b) = (1, 0).$$

$$\text{Hence } (1/a, -b/a) \text{ is the inverse of } (a, b)$$

$$\text{Hence } (G, *) \text{ is a group.}$$

19. In N we define $a * b = a$. Then $(N, *)$ is not a group.

Proof:

Clearly $*$ is an associative binary operation on N . However, there is no element $e \in N$ such that $e * a = a$ for all $a \in N$. Hence there is no identity element in $(N, *)$.

Hence $(N, *)$ is not a group.

In the group $(Z, +)$, the binary operation is commutative whereas in the group of 2×2 non-singular matrices, the matrix multiplication is not commutative.

Definition: 2.0.3

A group G is said to be **abelian** if $ab = ba$ for all $a, b \in G$. A group which is not abelian is called a **non-abelian** group.

Examples:

1. Z, Q, R and C under usual addition are abelian groups.

Space for Hints

2. $(\wp(S), \Delta)$ is an abelian group since $A \Delta B = B \Delta A$ for all $A, B \in \wp(S)$.
3. Let $B(R)$ denote the set of all bijections from R . Then $B(R)$ is a group under the composition of functions. This group is non-abelian.

For, consider

$f : R \rightarrow R$ given by $f(x) = x + 3$ and

$g : R \rightarrow R$ given by $g(x) = 2x$.

Clearly f and g are bijections.

Now $(f \circ g)(x) = f[g(x)] = f(2x) = 2x + 3$ and

$(g \circ f)(x) = g[f(x)] = g(x + 3) = 2x + 6$.

Hence $f \circ g \neq g \circ f$.

Hence $B(R)$ is non-abelian.

4. $(Z_n, \oplus)$ is an abelian group.
5. Consider the group given in example 18 of 2.0.1 Here
 $(2,3) * (4,5) = (8,17)$ and $(4,5) * (2,3) = (8,13)$. Thus
 $(2,3) * (4,5) \neq (4,5) * (2,3)$ so that this group is non-abelian.

2.0.4 Elementary Properties of a Group:

Theorem: 2.1

Let G be a group. Then

- (i) Identity element of G is unique.
- (ii) For any $a \in G$, the inverse of a is unique.

Proof:

(i) Let e and e' be two identity elements of G . Then $ee' = e'$ (since e' is an identity). Hence $e = e'$.

(ii) Let a' and a'' be two inverse of a . Hence $aa' = a'a = e$ and $aa'' = a''a = e$.

$\therefore a' = a'e = a'(aa'') = (a'a)a'' = ea'' = a''$.

Note:

We denote the inverse of a by a^{-1} .

Theorem: 2.2

In a group the left and right cancellation laws holds (ie)
 $ab = ac \Rightarrow b = c$ and $ba = ca \Rightarrow b = c$.

Proof:

$$\begin{aligned}
 ab = ac & \Rightarrow a^{-1}(ab) = a^{-1}(ac) \\
 & \Rightarrow (a^{-1}a)b = (a^{-1}a)c \\
 & \Rightarrow eb = ec \\
 & \Rightarrow b = c.
 \end{aligned}$$

Similarly we can prove that $ba = ca \Rightarrow b = c$.

Theorem: 2.3

Let G be a group and $a, b \in G$. Then the equations $ax = b$ and $ya = b$ have unique solutions for x and y in G .

Proof:

Consider $a^{-1}b \in G$

Then $a(a^{-1}b) = (aa^{-1})b = eb = b$.

Hence $a^{-1}b$ is a solution of $ax = b$.

Now, to prove the uniqueness, let x_1 and x_2 be two solutions of $ax = b$. Then $ax_1 = b$ and $ax_2 = b$.

$\therefore ax_1 = ax_2$ which implies $x_1 = x_2$.

$\therefore x = a^{-1}b$ is the unique solution for $ax = b$.

Similarly we can prove that $y = ba^{-1}$ is the unique solution of the equation $ya = b$.

Theorem: 2.4

Let G be a group. Let $a, b \in G$. Then $(ab)^{-1} = b^{-1}a^{-1}$ and $(a^{-1})^{-1} = a$.

Proof:

$$(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aea^{-1} = aa^{-1} = e.$$

Similarly $(b^{-1}a^{-1})(ab) = e$.

Hence $(ab)^{-1} = b^{-1}a^{-1}$.

Proof of the second part is obvious.

Corollary:

If $a_1, a_2, \dots, a_n \in G$ then $(a_1, a_2, \dots, a_n)^{-1} = a_n^{-1}a_{n-1}^{-1} \dots a_1^{-1}$

Definition: 2.1.4

Let G be a group and $a \in G$. For any positive integer n , we define $a^n = aa \dots a$ (a written n times).

$$\begin{aligned} \text{Clearly } (a^n)^{-1} &= (aa \dots a)^{-1} \\ &= (a^{-1}a^{-1} \dots a^{-1}) = (a^{-1})^n. \end{aligned}$$

We now define $a^{-n} = (a^{-1})^n = (a^n)^{-1}$.

Finally we define $a^0 = e$. Thus a^n is defined for all integers n .

Note:

When the binary operation on G is "+", we denote $a+a+\dots+a$ (a written n times) as na .

Theorem: 2.5

- (i) $a^m a^n = a^{m+n}, m, n \in \mathbb{Z}$.
- (ii) $(a^m)^n = a^{mn}, m, n \in \mathbb{Z}$.

Note:

In the additive notation, the above results take the form $ma + na = (m+n)a$ and $n(ma) = (nm)a$.

Problems:

Problem: 1

Show that in a group G , $x^2 = x$ if and only if $x = e$.

Solution:

Clearly $e^2 = ee = e$.

Conversely, let $x^2 = x$.

Then $xx = xe$. Hence by cancellation law $x = e$.

Note:

An element $a \in G$ is called **idempotent** if $a^2 = a$. Thus we have shown that in a group G , the identity element is the only idempotent element.

Problem: 2

In an abelian group $(ab)^2 = a^2b^2$.

Solution:

$$(ab)^2 = (ab)(ab) = a(ba)b = a(ab)b = (aa)(bb) = a^2b^2.$$

Note:

In general for any positive integer n , $(ab)^n = a^n b^n$ (prove by using induction).

Problem: 3

Let G be a group such that $a^2 = e$ for all $a \in G$. Then G is abelian.

Solution:

$$a^2 = e \Rightarrow aa = e \Rightarrow a = a^{-1}.$$

$$\text{Now, } ab = (ab)^{-1} = b^{-1}a^{-1} = ba.$$

Hence G is abelian.

Problem: 4

Let G be a group in which $(ab)^m = a^m b^m$ for three consecutive integers and for all $a, b \in G$. Then G is abelian.

Solution:

Let $a, b \in G$.

$$\text{Let } (ab)^m = a^m b^m : (ab)^{m+1} = a^{m+1} b^{m+1} \text{ and } (ab)^{m+2} = a^{m+2} b^{m+2}.$$

$$\text{Now, } (ab)^{m+1} = a^{m+1} b^{m+1}$$

$$\Rightarrow (ab)^m (ab) = (a^m a)(b^m b)$$

$$\Rightarrow (a^m b^m)(ab) = (a^m a)(b^m b)$$

$$\Rightarrow b^m a = ab^m \text{ (by cancellation law)} \quad (1)$$

$$\text{Similarly } (ab)^{m+2} = a^{m+2} b^{m+2}$$

$$\Rightarrow b^{m+1} a = ab^{m+1}$$

$$\Rightarrow b^m ba = ab^m b$$

$$\Rightarrow b^m ba = b^m ab \text{ (by (1))}$$

$$\Rightarrow ba = ab \text{ (by cancellation law)}$$

$\therefore G$ is abelian.

Problem: 5

Let $(H, \cdot)$ and $(K, *)$ be groups. We define a binary operation $\square$ on $H \times K$ by $(h_1, k_1) \square (h_2, k_2) = (h_1 h_2, k_1 * k_2)$.

Then $H \times K$ is a group.

Note:

$H \times K$ is called the **direct product** of H and K .

Solution:

First we shall prove that $\square$ is associative.

Let $(h_1, k_1), (h_2, k_2), (h_3, k_3) \in H \times K$.

$$\begin{aligned} & [(h_1, k_1) \square (h_2, k_2)] \square (h_3, k_3) \\ &= (h_1 h_2, k_1 * k_2) \square (h_3, k_3) \\ &= ((h_1 h_2) h_3, (k_1 * k_2) * k_3) \\ &= (h_1 (h_2 h_3), k_1 * (k_2 * k_3)) \\ &= (h_1, k_1) \square (h_2 h_3, k_2 * k_3) \\ &= (h_1, k_1) \square [(h_2, k_2) \square (h_3, k_3)]. \end{aligned}$$

Let e, e_1 be the identities of the groups H and K respectively. Clearly (e, e_1) is the identity element in $H \times K$. Also (h^{-1}, k^{-1}) is the inverse of (h, k) .

Hence $H \times K$ is a group.

2.1 Subgroups:

Definition: 2.1.1

Let G be a set with a binary operation $*$ defined on it. Let $S \subseteq G$. If for each $a, b \in S, a * b$ (computed in G) is in S , we say that S is **closed** with respect to the **binary operation** “ $*$ ”

Examples:

1. $(\mathbb{Z}, +)$ is a group. The set E of all even integers is closed under $+$ and further $(E, +)$ is itself a group.
2. The set of G of all non-singular 2×2 matrices from a group under matrix multiplication. Let H be the set of all matrices of the form $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$. H is subset of G . Also H itself is a group under matrix multiplication.

Definition: 2.1.2

A subset H of group G is called a **subgroup** of G if H forms a group with respect to the binary operation in G .

Examples:

Space for Hints

1. Let G be any group. Then $\{e\}$ and G are subgroup of G . They are called **improper subgroups of G** .
2. $(\mathbb{Q}, +)$ is a subgroup of $(\mathbb{R}, +)$ and $(\mathbb{R}, +)$ is a subgroup of $(\mathbb{C}, +)$.
3. In $(\mathbb{Z}_8, \oplus)$, let $H_1 = \{0, 4\}$ and $H_2 = \{0, 2, 4, 6\}$. The Cayley tables for

H_1 and H_2 are given by

$\oplus$	0	4
0	0	4
4	4	0

$\oplus$	0	2	4	6
0	0	2	4	6
2	2	4	6	0
4	4	6	0	2
6	6	0	2	4

It is easily seen that H_1 and H_2 are closed under $\oplus$ and $(H_1, \oplus)$ and $(H_2, \oplus)$ are groups. Hence H_1 and H_2 are subgroups of $\mathbb{Z}_8$.

4. $\{1, -1\}$ is a subgroup of $(\mathbb{R}^*, \cdot)$.
5. $\{1, i, -1, -i\}$ is a subgroup of $(\mathbb{C}^*, \cdot)$.
6. For any integer n we define $n\mathbb{Z} = \{nx / x \in \mathbb{Z}\}$. Then $(n\mathbb{Z}, +)$ is a subgroup of $(\mathbb{Z}, +)$.

For, let $a, b \in n\mathbb{Z}$. Then $a = nx$ and $b = ny$ where $x, y \in \mathbb{Z}$.

Hence $a + b = n(x + y) \in n\mathbb{Z}$: Hence $n\mathbb{Z}$ is closed under $+$.

$0 \in n\mathbb{Z}$ is the identity element.

Inverse of nx is $-nx = n(-x) \in n\mathbb{Z}$.

Hence $(n\mathbb{Z}, +)$ is a group.

7. In the symmetric group S_3 , $H_1 = \{e, p_1, p_2\}$; $H_2 = \{e, p_3\}$; $H_3 = \{e, p_4\}$; and $H_4 = \{e, p_5\}$ are subgroups.

8. A_n is a subgroup of S_n .

9. The set of permutations $\{e, p_1, p_2, p_3\}$ where

$$e = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}; \quad p_1 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix};$$

$$p_2 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}; \quad p_3 = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}$$

is a subgroups of S_4 .

Note:

In all the above the examples we see that the identity element in the subgroup is the same as the identity element of the group.

Theorem: 2.6

Let H be a subgroup of G . Then

- (a) The identity element of H is the same as that of G .
- (b) For each $a \in H$ the inverse of a in H is the same as the inverse of a in G .

Proof:

- (a) Let e and e' be the identities of G and H respectively.

Let $a \in H$. Now,

$$\begin{aligned} e'a &= a \text{ (since } e' \text{ is the identity of } H) \\ &= ea \text{ (since } e \text{ is the identity of } G \text{ and } a \in G) \end{aligned}$$

$$\therefore e'a = ea.$$

$$\therefore e' = e \text{ (by cancellation law).}$$

- (b) Let a' and a'' be the inverse of a in G and H respectively. Since by (a), G and H have the same identity element e , we have $a'a = e = a''a$. Hence by cancellation law $a' = a''$.

Theorem: 2.7

A subset H of a group G is a subgroup of G iff

- (i) It is closed under the binary operation in G .
- (ii) The identity e of G is in H .
- (iii) $a \in H \Rightarrow a^{-1} \in H$.

Proof:

Let H be a subgroup of G . The result follows immediately from theorem 2.6.

Conversely let H be a subset of G satisfying conditions (i), (ii) and (iii). Then, obviously H itself is a group with respect to the binary operation in G .

Therefore H is a subgroup of G .

Theorem: 2.8

A non - empty subset H of a group G is a subgroup of G iff

$$a, b \in H \Rightarrow ab^{-1} \in H.$$

Proof:

Let H be a subgroup of G . Then

$$a, b \in H \Rightarrow a, b^{-1} \in H \Rightarrow ab^{-1} \in H.$$

Conversely let H be a non - empty subset of G such that

$$a, b \in H \Rightarrow ab^{-1} \in H.$$

Since $H \neq \Phi$, there exists an element $a \in H$.

Hence $aa^{-1} \in H$. Thus $e \in H$

Also since $e, a \in H, ea^{-1} \in H$. Hence $a^{-1} \in H$. Now, let $a, b \in H$.

Then $a, b^{-1} \in H$.

Hence $a(b^{-1})^{-1} = ab \in H$. Thus H is closed under the binary operation in G .

Hence by theorem 2.7 H is a subgroup of G .

Note:

If the operation is $+$ then H is a subgroup of G iff $a, b \in H \Rightarrow a - b \in H$.

Theorem: 2.9

Let H be a non - empty finite subset of G . If H is closed under the operation in G then H is a subgroup of G .

Proof:

Let $a \in H$.

Since H is closed $a, a^2, a^3, \dots, a^n, \dots$ are all elements of H .

But since H is finite the elements $a, a^2, a^3, \dots$ cannot all be distinct.

Hence let $a^r = a^s, r < s$. Then $a^{s-r} = e \in H$.

Now, let $a \in H$. We have proved that $a^n = e$ for some n . Hence $aa^{n-1} = e$. Hence $a^{-1} = a^{n-1} \in H$.

Thus H is a subgroup of G .

Note:

The above theorem is not true if H is infinite. For example, N is an infinite subset of $(\mathbb{Z}, +)$ and N is closed under addition. However N is

not a subgroup of $(\mathbb{Z}, +)$.

Theorem: 2.10

If H and K are subgroups of a group G then $H \cap K$ is also a subgroup of G .

Proof:

Clearly $e \in H \cap K$ and hence $H \cap K$ is non - empty. Now let $a, b \in H \cap K$. Then $a, b \in H$ and $a, b \in K$. Since H and K are subgroups of G , $ab^{-1} \in H$ and $ab^{-1} \in K$.

$\therefore ab^{-1} \in H \cap K$. Hence by theorem 2.8, $H \cap K$ is a subgroup of G .

Note:

1. It can be similarly proved that the intersection of any number of subgroups of G is again a subgroup of G .
2. The union of two subgroups of a group need not be a subgroup. For example, $2\mathbb{Z}$ and $3\mathbb{Z}$ are subgroups of $(\mathbb{Z}, +)$ but $2\mathbb{Z} \cup 3\mathbb{Z}$ is not a subgroup of $\mathbb{Z}$ since $3, 2 \in 2\mathbb{Z} \cup 3\mathbb{Z}$ but $3 + 2 = 5 \notin 2\mathbb{Z} \cup 3\mathbb{Z}$.

Theorem: 2.11

The union of two subgroups of a group G is a subgroup iff one is contained in the other.

Proof:

Let H and K be two subgroups of G such that one is contained in the other. Hence either $H \subseteq K$ or $K \subseteq H$.

$\therefore H \cup K = K$ or $H \cup K = H$. Hence $H \cup K$ is a subgroup of G .

Conversely, suppose $H \cup K$ is a subgroup of G . We claim that $H \subseteq K$ or $K \subseteq H$.

Suppose that H is not contained in K and K is not contained in H . Then there exist elements a, b such that

$$a \in H \quad \text{and} \quad a \notin K \quad (1)$$

$$b \in K \quad \text{and} \quad b \notin H \quad (2)$$

Clearly, $a, b \in H \cup K$. Since $H \cup K$ is a subgroup of G , $ab \in H \cup K$. Hence $ab \in H$ or $ab \in K$.

Case: (i)

Let $ab \in H$. Since $a \in H, a^{-1} \in H$.

Hence $a^{-1}(ab) = b \in H$ which is a contradiction to (2).

Case: (ii)

Let $ab \in K$. Since $b \in K, b^{-1} \in K$.

Hence $(ab)b^{-1} = a \in K$ which is a contradiction to (1). Hence our assumption that H is not contained in K and K is not contained in H is false.

$\therefore H \subseteq K$ or $K \subseteq H$.

Definition: 2.1.3

Let A and B be two subsets of a group G . We define $AB = \{ab / a \in A, b \in B\}$.

Note:

If A and B are two subgroups of G . AB need not be a subgroup of G .

In S_3 , consider $A = \{e, p_3\}$ and $B = \{e, p_4\}$. Clearly A and B are subgroups of S_3 .

Also $AB = \{ee, ep_4, ep_3, p_3p_4\} = \{e, p_4, p_3, p_2\}$.

Now $p_4p_2 = p_5 \notin AB$.

Hence AB is not a subgroup of S_3 .

Theorem: 2.12

Let A and B be two subgroups of a group G . Then AB is a subgroup of G iff $AB = BA$.

Proof:

Let AB be a subgroup of G .

We claim that $AB = BA$.

Let $x \in AB$. Since AB is a subgroup of G , $x^{-1} \in AB$.

Let $x^{-1} = ab$ where $a \in A$ and $b \in B$.

$\therefore x = (ab)^{-1} = b^{-1}a^{-1}$.

Since A and B are subgroups of G , $a^{-1} \in A$ and $b^{-1} \in B$.

$$\therefore x \in BA. \text{ Hence } AB \subseteq BA \quad (1)$$

Now, let $x \in BA$. Then $x = ba$ where $b \in B$ and $a \in A$.

$$\therefore x^{-1} = (ab)^{-1} = a^{-1}b^{-1} \in AB.$$

Now, since AB is a subgroup and $x^{-1} \in AB$, we have $x \in AB$.

$$\therefore BA \subseteq AB \quad (2)$$

From (1) and (2) we get $AB = BA$.

Conversely, let $AB = BA$. We claim that AB is a subgroup of G . Clearly $e \in AB$ and hence AB is non-empty. Now let $x, y \in AB$. Then $x = a_1b_1$ and $y = a_2b_2$ where $a_1, a_2 \in A$ and $b_1, b_2 \in B$.

$$\therefore xy^{-1} = (a_1b_1)(a_2b_2)^{-1} = a_1b_1b_2^{-1}a_2^{-1}.$$

Now, $b_2^{-1}a_2^{-1} \in BA$. Since $BA = AB$, $b_2^{-1}a_2^{-1} \in AB$.

$$\therefore b_2^{-1}a_2^{-1} = a_3b_3 \text{ where } a_3 \in A \text{ and } b_3 \in B.$$

$$\therefore xy^{-1} = a_1b_1a_3b_3.$$

Now $b_1a_3 \in BA$. Since $BA = AB$, $b_1a_3 \in AB$.

$$\therefore b_1a_3 = a_4b_4 \text{ where } a_4 \in A \text{ and } b_4 \in B.$$

$$\therefore xy^{-1} = a_1(a_4b_4)b_3 = (a_1a_4)(b_4b_3) \in AB.$$

$\therefore AB$ is subgroup of G .

Corollary:

If A and B are subgroups of an abelian group G , then AB is a subgroup of G .

Proof:

Let $x \in AB$. Then $x = ab$ where $a \in A$ and $b \in B$. Since G is abelian, $ab = ba$.

$$\therefore x \in BA. \text{ Hence } AB \subseteq BA.$$

Similarly $BA \subseteq AB$.

$$\therefore AB = BA.$$

Hence AB is a subgroup of G .

Problems:

Problem: 1

Let $a \in \mathbb{R}^*$. Let $H = \{a^n / n \in \mathbb{Z}\}$. Then H is a subgroup of $\mathbb{R}^*$.

Solution:

Clearly H is non – empty.

Now, let $x, y \in H$.

Then $x = a^s$ and $y = a^t$ where $s, t \in \mathbb{Z}$.

$$\therefore xy^{-1} = a^s (a^t)^{-1} = a^{s-t} \in H.$$

Hence H is a subgroup of R^* .

Problem: 2

Let H denote the set of all permutations in S_n fixing the symbol

1. Then H is a subgroup of S_n :

Solution:

Clearly $e \in H$ and hence H is non – empty. Let $\alpha, \beta \in H$. Then α and β fix the symbol 1. Now β fixes the symbol 1 $\Rightarrow \beta^{-1}$ fixes the symbol 1. Hence $\alpha\beta^{-1}$ fixes the symbol 1. Hence $\alpha\beta^{-1} \in H$.

Thus H is a subgroup of S_n .

Problem: 3

Let G be the set of all 2×2 matrices with entries from $\mathbb{R}$. Then G is a group under matrix addition.

Let $H = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$. Then H is a subgroup of G .

Solution:

Let $A, B \in H$.

$$\text{Then } A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} \text{ and } B = \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}$$

$$\text{Now, } A - B = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} - \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix} = \begin{pmatrix} a-c & 0 \\ 0 & b-d \end{pmatrix} \in H.$$

Hence H is a subgroup of G .

Problem: 4

Let G be a group.

Let $H = \{a \mid a \in G \text{ and } ax = xa \text{ for all } x \in G\}$. (ie) H is the set of all elements which commute with every other element. Show that H is a subgroup of G .

Clearly $ex = xe = x$ for all $x \in G$.

Hence $e \in H$, so that H is non empty.

Now, let $a, b \in H$.

Then $ax = xa$ and $bx = xb$ for all $x \in G$.

Now,

$$\begin{aligned} bx = xb &\Rightarrow b^{-1}(bx)b^{-1} = b^{-1}(xb)b^{-1} \\ &\Rightarrow (b^{-1}b)xb^{-1} = b^{-1}x(bb^{-1}) \\ &\Rightarrow exb^{-1} = b^{-1}xe \\ &\Rightarrow xb^{-1} = b^{-1}x. \end{aligned} \quad (1)$$

$$\begin{aligned} \therefore (ab^{-1})x &= a(b^{-1}x) \\ &= a(xb^{-1}) \quad (\text{by (1)}) \\ &= (ax)b^{-1} \\ &= (xa)b^{-1} \quad (\text{since } ax = xa) \\ &= x(ab^{-1}). \end{aligned}$$

Thus ab^{-1} commutes with every element of G .

$\therefore ab^{-1} \in H$ and hence H is a subgroup of G .

Note:

The above subgroup of G is called the **center of G** and is denoted by $Z(G)$.

Problem: 5

Let G be a group and let a be a fixed element of G .

Let $H_a = \{x / x \in G \text{ and } ax = xa\}$

(ie) H_a is the set of all elements in G which commute with a .

Show that H_a is a subgroup of G .

Solution:

Clearly $ea = ae = a$.

Hence $e \in H_a$ so that H_a is non – empty.

Now, let $x, y \in H_a$.

Then $ax = xa$ and $ay = ya$.

Now, $ay = ya \Rightarrow y^{-1}a = ay^{-1}$. (as in the previous problem) (1)

$$\begin{aligned} \text{Hence } a(xy^{-1}) &= (ax)y^{-1} \\ &= (xa)y^{-1} \quad (\text{since } ax = xa) \\ &= x(ay^{-1}) \\ &= x(y^{-1}a) \quad (\text{by (1)}) \\ &= (xy^{-1})a. \end{aligned}$$

Hence xy^{-1} commutes with a .

$\therefore xy^{-1} \in H_a$ and hence H_a is a subgroup of G .

Note:

H_a is called the **normaliser** of a in G .

Exercises:

1. Show that $\{a + bi / a, b \in \mathbb{Z}\}$ is a subgroup of $(\mathbb{C}, +)$.
2. Determine which of the following are subgroups of $(\mathbb{C}, +)$.
 - (a) $\mathbb{R}$
 - (b) $\{a + b\sqrt{-5} / a, b \in \mathbb{N}\}$
 - (c) $\{z / |z| = a\}$
 - (d) $\{z / \text{real part of } z \text{ is } 0\}$
 - (e) $\{1, i, -1, -i\}$.
3. Let G_1 and G_2 be two groups. Let e_1 and e_2 be the identity elements of G_1 and G_2 respectively. Let $G_1 \times G_2$ be the direct product of these groups. Let $H = \{(e_1, y) / y \in G_2\}$ and $K = \{(x, e_2) / x \in G_1\}$. Show that H and K are subgroups of $G_1 \times G_2$.
4. Let G be a group and let H be the centre of G . Show that $H = G$ iff G is abelian.
5. Show that the centre of S_3 is $\{e\}$.
(Hint: For each $a \in S_3$ and $a \neq e$, find another element $b \in S_3$ such that $ab \neq ba$).
6. Show that a proper subgroup of a non-abelian group can be abelian.

(Hint: Consider any proper subgroup of S_3).

7. Show that any subgroup of an abelian group is abelian.
8. Let S and N be subgroups of G such that $S \cap N = \{e\}$ and $S \cup N = G$.
Prove that either $S = G$ or $N = G$.

9. Find as many subgroups as you can in

- (a) V_4
- (b) The group of symmetries of a square.
- (c) Z_6
- (d) Z

2.2 Permutation Groups:

The set of all bijections $B(A)$ from A to itself is a group under the composition of functions. In this section we make detailed study of this group when A is finite.

Definition: 2.2.1

Let A be a finite set. A bijection from A to itself is called a **permutation** of A .

For example, if $A = \{1, 2, 3, 4\}$ $f: A \rightarrow A$ given by $f(1) = 2$, $f(2) = 1$, $f(3) = 4$ and $f(4) = 3$ is permutation of A . We shall write this

$$\text{permutation } \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}.$$

An element in the bottom row is the image of the element just above it in the upper row.

Note:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = \begin{pmatrix} 4 & 3 & 1 & 2 \\ 3 & 4 & 2 & 1 \end{pmatrix}$$

Hence any rearrangement of columns in a permutation is immaterial.

Definition: 2.2.2

Let A be a finite set containing n element. The set of all permutations of A is clearly group under the composition of functions. This group is called the **symmetric group** of degree n and denoted by S_n .

Example:

Let $A = \{1, 2, 3\}$. Then S_3 consists of

$$e = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}; p_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}; p_2 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}; p_3 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix};$$

$$p_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}; p_5 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

In this group, e is the identity element. We now compute the product

$p_1 p_2$.

$$\begin{array}{ccc} & 1 & 2 & 3 \\ p_1: & \downarrow & \downarrow & \downarrow \\ & 2 & 3 & 1 \\ p_2: & \downarrow & \downarrow & \downarrow \\ & 1 & 2 & 3 \end{array} \quad \text{Hence } p_1 p_2: \begin{array}{ccc} & 1 & 2 & 3 \\ & \downarrow & \downarrow & \downarrow \\ & 1 & 2 & 3 \end{array}$$

so that $p_1 p_2 = e$.

$$\text{Now, } p_1 p_4 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = p_5.$$

Similarly we can compute all the other products and Cayley table for this group is given by

$\circ$	e	p_1	p_2	p_3	p_4	p_5
e	e	p_1	p_2	p_3	p_4	p_5
p_1	p_1	p_2	e	p_4	p_5	p_3
p_2	p_2	e	p_1	p_5	p_3	p_4
p_3	p_3	p_5	p_4	e	p_2	p_1
p_4	p_4	p_3	p_5	p_1	e	p_2
p_5	p_5	p_4	p_3	p_2	p_1	e

Thus S_3 is a group containing $3! = 6$ elements.

Remarks:

1. In definition 1.2.15 we have defined the composition $g \circ f$ of two functions f and g by $(g \circ f)(x) = g[f(x)]$.

Hence to find the image of any element x under $g \circ f$, we first apply f and then g . However in forming the product of two permutations p_1 and p_2 we adopt a different convention. To find the image of x under the product $p_1 p_2$, we first apply p_1 and then p_2 .

2. In S_3 , $p_1 p_2 = p_2 p_1 = e$ so that the inverse of p_1 is p_2 . In general the inverse of a permutation can be obtained by interchanging the rows of the permutation.

For example, if $p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 2 & 5 & 1 \end{pmatrix}$ then the inverse of p is the

permutation given by $p^{-1} = \begin{pmatrix} 3 & 4 & 2 & 5 & 1 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 2 & 4 \end{pmatrix}$

3. In S_3 , $p_1 p_4 = p_5$ and $p_4 p_1 = p_3$. Hence $p_1 p_4 \neq p_4 p_1$ so that S_3 is non-abelian.

4. The symmetric group S_n contains $n!$ elements, for let $A = \{1, 2, \dots, n\}$. Any permutation on A is given by specifying the image of each element. The image of 1 can be chosen in n different ways. Since the image of two is different from the image of 1, it can be chosen in $(n - 1)$ different ways and so on.

Hence the number of permutations of A is $n(n - 1) \dots 2 \cdot 1 = n!$ so that the number of elements in S_n is $n!$.

Definition: 2.2.3

Let G be a finite group. Then the number of elements in G is called the order of G and is denoted by $|G|$ or $\circ(G)$.

Exercises:

1. Compute $\alpha\beta, \beta\alpha$ and α^{-1} if

(a) $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 1 & 3 \end{pmatrix}$

$\beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \end{pmatrix}$

(b) $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 4 & 6 & 5 \end{pmatrix}$

$\beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 4 & 5 & 6 & 1 \end{pmatrix}$

(c) $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 4 & 7 & 2 & 5 & 6 & 1 \end{pmatrix}$

$$\beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 7 & 6 & 2 & 3 & 4 & 5 & 1 \end{pmatrix}$$

2. Show that the permutation $e = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}$;

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}; \beta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \end{pmatrix}; \gamma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} \text{ from a}$$

group and construct its Cayley table.

3. From the Cayley table for the group S_2 .

$$\text{Consider the permutation } \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}$$

In this permutation $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 5 \rightarrow 1$. Thus the permutation maps the symbols in a cyclic order. Now consider the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 5 & 2 \end{pmatrix}. \text{ This permutation fixes the symbol 1 and maps}$$

the remaining symbols in a cyclic order.

Definition: 2.2.4

Let p be a permutation on

$A = \{1, 2, \dots, n\}$. p is called a **cycle of length r** if there exist

distinct symbols $a_1, a_2, \dots, a_r$ such that

$p(a_1) = a_2, p(a_2) = a_3, \dots, p(a_{r-1}) = a_r$, and $p(a_r) = a_1$, and $p(b) = b$,

for all

$$b \in A - \{a_1, a_2, \dots, a_r\}.$$

This cycle is represented by the symbol $(a_1, a_2, \dots, a_r)$. Thus under the cycle $(a_1, a_2, \dots, a_r)$ each symbol is mapped onto the following symbol except the last one which is mapped onto the first symbol and all the other symbols not in the cycle are fixed.

Example:

Let $A = \{1, 2, 3, 4, 5\}$. Consider the cycle of length 4 given by $p = (2451)$.

$$\text{Then } p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 3 & 5 & 1 \end{pmatrix}.$$

$$\text{Obviously } (2451) = (4512) = (5124) = (1245).$$

Note:

Since cycles are special types of permutations, they can be multiplied in the usual way. The product of cycles need not be a cycle.

For example, let $p_1 = (234)$ and $p_2 = (1,5)$. Then

$$p_1 p_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 3 & 4 & 2 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 3 & 4 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 4 & 2 & 1 \end{pmatrix} \text{ which is not}$$

a cycle.

Definition: 2.2.5

Two cycles are said to be **disjoint** if they have no symbols in common.

For example $(2\ 1\ 5)$ and $(3\ 4)$ are disjoint cycles.

Note:

If p_1 and p_2 are disjoint cycles the symbols which are moved by p_1 are fixed by p_2 and vice versa. Hence multiplication of disjoint cycles is commutative.

Examples:

1. Consider the permutation $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 1 & 3 & 5 & 6 & 7 & 4 \end{pmatrix}$. We shall write

this permutation as a product of disjoint cycles. First of all 1 is moved to 2 and then 2 is moved to 1 thus giving the cycle $(1\ 2)$. The element 3 is left fixed. Again starting with 4, 4 is moved to 5, 5 is moved to 6, 6 is moved to 7 and 7 is moved to 4, thus giving the cycle $(4\ 5\ 6\ 7)$. Thus

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 1 & 3 & 5 & 6 & 7 & 4 \end{pmatrix} = (12)(4567) = (4567)(12).$$

2. Consider the permutation $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 3 & 7 & 5 & 4 & 1 & 6 \end{pmatrix} \in S_7$.

Starting with 1 we get the cycle $(1\ 2\ 3\ 7\ 6)$. The elements 4, 5 do not

appear in it. Starting with 4 we get the cycle (4 5). Each element of the set $\{1, 2, \dots, 7\}$ occurs in one of these two cycles.

Thus $\alpha = (12376)(45)$.

3. Consider the permutation $\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 4 & 6 & 1 & 3 & 2 & 5 \end{pmatrix}$. Clearly $\alpha = (143)(265)$.

Theorem: 2.13

Any permutation can be expressed as a product of disjoint cycles.

Proof:

Let p be a given permutation of the set $S = \{1, 2, \dots, n\}$. Let us start with any symbol $a_1 \in S$. Let $p(a_1) = a_2, p(a_2) = a_3, \dots$. Since S is finite, these symbols cannot all be distinct and hence there exists a least positive integer r such that $1 \leq r \leq n$ and $p(a_r) = a_1$.

Let $c = (a_1, a_2, \dots, a_r)$. If $r = n$ then $p = c$ so that p is a cycle. If $r < n$, let b_1 be a symbol in S such that $b_1 \notin (a_1, a_2, \dots, a_r)$. Starting with b_1 we can construct the cycle $d = (b_1, b_2, \dots, b_s)$ as before. Clearly the cycles c and d are disjoint. If $r + s = n$ then $p = cd$. If $r + s < n$ we repeat the above process to obtain more cycles until all the symbols appear in one of the cycles. Thus we get a decomposition of p into disjoint cycles.

Exercise:

Express the following permutations as a product of disjoint cycles.

(a) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 1 & 3 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix}$

(c) $(1\ 2\ 3\ 4)(3\ 4\ 5)$

(d) $(1\ 3)(3\ 4)(4\ 5)$

(e) $(1\ 2\ 3)(1\ 6\ 5\ 4\ 3)$

Space for Hints

$$(f) \quad (4 \ 2 \ 1 \ 5) (3 \ 4 \ 2 \ 6) (5 \ 6 \ 7 \ 1)$$

$$(g) \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 4 & 2 & 6 & 1 \end{pmatrix}$$

$$(h) \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 4 & 3 & 7 & 2 & 1 & 6 \end{pmatrix}$$

Answers:

$$(a) \quad (14) (35)$$

$$(b) \quad (135) (24)$$

$$(c) \quad (124) (35)$$

$$(d) \quad (1543)$$

$$(e) \quad (12) (3654)$$

$$(f) \quad (16347) (25)$$

$$(g) \quad (134256)$$

$$(h) \quad (152476).$$

Note:

The decomposition of a permutation into disjoint cycles is unique except for the order of the factors.

Definition: 2.2.6

A cycle of length two is called a **transposition**. Thus a transposition $(a_1 a_2)$ interchanges the symbols a_1 and a_2 and leaves all the other elements fixed.

Theorem: 2.14

Any permutation can be expressed as a product of transpositions.

Proof:

Since any permutation is a product of disjoint cycles it is enough if we prove that each cycle is a product of transpositions.

Hence let $c = (a_1 a_2, \dots, a_r)$ be a cycle.

Clearly $(a_1 a_2, \dots, a_r) = (a_1 a_2)(a_1 a_3) \dots (a_1 a_r)$. This proves the theorem.

Examples:

$$1. \quad \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 2 & 1 \end{pmatrix} = (1245) = (12)(14)(15).$$

Also $(1235) = (2451) = (24)(25)(21)$.

Thus the representation of a permutation as a product of transpositions is not unique.

$$2. (1345)(26) = (13)(14)(15)(26) = (13)(12)(12)(14)(15)(26).$$

Thus in the representation of a permutation as a product of transpositions one can always insert $(ab)(ab)$ in any place since $(ab)(ab)$ is the identity permutation.

Theorem: 2.15

If a permutation $p \in S_n$ is a product of r transpositions and also a product of s transpositions then either r and s are both even or both odd.

Proof:

Let $p = t_1 t_2 \dots t_r = t'_1 t'_2 \dots t'_s$ where t_i, t'_i are transpositions.

Now consider the polynomial in n variables $x_1, x_2, \dots, x_n$ given by

$$\Delta = (x_1 - x_2)(x_1 - x_3) \dots (x_1 - x_n)(x_2 - x_3)(x_2 - x_4) \dots (x_2 - x_n) \dots x(x_{n-1} - x_n) = \prod_{i < j} (x_i - x_j)$$

For any permutation $p \in S_n$ we define

$$p(\Delta) = \prod_{i < j} (x_{p(i)} - x_{p(j)}).$$

Consider the transposition $t = (ij)$. Then the factor $x_i - x_j$ in Δ becomes $x_j - x_i$. Any factor $(x_k - x_\ell)$ of Δ in which neither i nor j is equal to k or ℓ is unchanged. All other factors of Δ can be paired to form products of the form $\pm(x_i - x_k)(x_k - x_j)$, the sign being determined by the relative magnitudes of i, j and k . Since t interchanges x_i and x_j any such product is unchanged. Hence the effect of the transposition t on Δ is just to change the sign of Δ ie, $t(\Delta) = -\Delta$.

$$\therefore p(\Delta) = (t_1 t_2 \dots t_r)(\Delta) = (-1)^r \Delta.$$

$$\text{Also } p(\Delta) = (t'_1 t'_2 \dots t'_s)(\Delta) = (-1)^s \Delta.$$

$$\therefore (-1)^r = (-1)^s \Rightarrow r \text{ and } s \text{ are both even or both odd.}$$

Definition: 2.2.7

A permutation $p \in S_n$ is called **even** or **odd** according as p can be expressed as a product of an even number of transpositions or an odd

Space for Hints

number of transpositions respectively.

Examples:

1. Consider the permutation $p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 3 & 6 & 4 & 1 & 7 & 2 & 5 \end{pmatrix}$

$$p = (134)(26)(57) = (13)(14)(26)(57).$$

$\therefore$ p is a product of 4 transpositions.

Hence p is an even permutation.

2. Consider the permutation $p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 5 & 4 & 3 & 6 & 1 & 7 & 9 & 8 \end{pmatrix}$

$$p = (1256)(34)(89) = (12)(15)(16)(34)(89)$$

$\therefore$ p is a product of 5 transpositions.

Hence p is an odd permutation.

Exercises:

1. Determine which of the following permutations are odd and which of them are even.

(a) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 5 & 6 & 4 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 3 & 4 & 2 & 1 & 6 & 5 & 8 & 7 \end{pmatrix}$

(c) $(1234)(356)(67)$

(d) $(123)(45)(5672).$

2. Find all the even permutations in S_3 and show that they form a group.

3. For what values of m is a cycle of length m an even permutation?

Answer:

1.

(a) Even

(b) Odd

(c) Even

(d) Even

2. e, p_1, p_2

3. m is odd.

Theorem: 2.16.

- (i). The product of two even permutations is an even permutation.
- (ii). The product of two odd permutations is an even permutation.
- (iii). The product of an even permutation and an odd permutation is an odd permutation.
- (iv). The inverse of an even permutation is an even permutation.
- (v). The inverse of an odd permutation is an odd permutation.
- (vi). The identity permutation e is an even permutation.

Proof:

Let p_1, p_2 be two permutations. If p_1 is a product of r transpositions and p_2 is a product of s transpositions, then $p_1 p_2$ is a product of $r + s$ transpositions. Hence (i), (ii) and (iii) follow.

Now suppose that a permutation p is a product of r transpositions, say, $p = t_1, t_2, \dots, t_r$. Then

$$p^{-1} = (t_1, t_2, \dots, t_r)^{-1} = t_r^{-1} \dots t_2^{-1} t_1^{-1} = t_r \dots t_2 t_1$$

$\therefore p^{-1}$ is also a product of r transpositions.

This proves (iv) and (v).

Now, $e = (12)(12)$ and hence e is an even permutation which proves (vi).

Theorem: 2.17

Let A_n be the set of all even permutations in S_n . Then A_n is a group containing $\frac{n!}{2}$ permutations.

Proof:

From (i), (vi) and (iv) of theorem 2.16 we see that A_n is a group.

Now let B_n be the set of all odd permutations in S_n .

Define $f : A_n \rightarrow B_n$ by $f(p) = (12)p$

f is 1-1, for $f(p_1) = f(p_2) \Rightarrow (12)p_1 = (12)p_2 \Rightarrow p_1 = p_2$.

f is onto, for, if $\alpha \in B_n$ then $(12)\alpha \in A_n$ and

$$f[(12)\alpha] = (12)(12)\alpha = \alpha$$

Thus f is a bijection and hence the number of odd permutations in S_n = the number of even permutations in S_n . Since S_n contains $n!$

permutations, A_n has $\frac{n!}{2}$ elements.

Definition: 2.2.8

The group A_n of all even permutations in S_n is called the **alternating group on n symbols**.

Exercises:

1. Let G be a group of permutations. Show that either all the permutations in G are even or exactly half of them are even.
2. Let p be a permutation of a set A . Let $a \in A$ we say that p moves a if $p(a) \neq a$. How many elements are moved by a cycle of length r ?
3. Show that the set of all permutations in S_n fixing the symbol 1 is a group.
4. Write down all the permutations of the set $\{1, 2, 3, 4\}$ and determine which of them are even.
5. Determine which of the following statements are true and which of them are false.
 - (a) Every cycle is a permutation.
 - (b) Every permutation is a cycle.
 - (c) Product of two cycles is a cycle.
 - (d) Any transposition is an odd permutation.
 - (e) When $n \geq 3$, S_n is nonabelian.
 - (f) Any permutation can be expressed as a product of cycles.
 - (g) The set of all odd permutations in S_n is a group.
 - (h) Any finite group is abelian.

Answers:

2. r elements

5.

- (a) True
- (b) False
- (c) False
- (d) True
- (e) True
- (f) True

(g) False

(h) False

2.3 Cyclic Groups:

Definition: 2.3.1

Let G be a group. Let $a \in G$.

Then $H' = \{a^n / n \in \mathbb{Z}\}$ is a subgroup of G (verify). H is called the **cyclic subgroup of G generated by a** and is denoted by $\langle a \rangle$.

Examples:

1. In $(\mathbb{Z}, +)$, $\langle 2 \rangle = 2\mathbb{Z}$ which is the group of even integers.

2. In the group

$$G = (\mathbb{Z}_{12}, \oplus), \langle 3 \rangle = \{0, 3, 6, 9\}, \langle 5 \rangle = \{0, 5, 10, 3, 8, 1, 6, 11, 4, 9, 2, 7\} = \mathbb{Z}_{12}.$$

3. In the group $G = \{1, i, -1, -i\}$, $\langle i \rangle = \{i, i^2, i^3, \dots\} = \{i, -1, -i, 1\} = G$.

Definition: 2.3.2

Let G be a group and let $a \in G$. a is called a **generator** of G if $\langle a \rangle = G$.

A group G is **cyclic** if there exists an element $a \in G$ such that $\langle a \rangle = G$.

Note:

If G is a cyclic group generated by an element a , then every element of G is of the form a^n for some $n \in \mathbb{Z}$.

Examples:

1. $(\mathbb{Z}, +)$ is a cyclic group. 1 is a generator of this group. -1 is also a generator of this group. Thus a cyclic group can have more than one generator.

2. $(n\mathbb{Z}, +)$ is a cyclic group, n and $-n$ are generator on this group.

3. $(\mathbb{Z}_8, \oplus)$ is a cyclic group. 1, 3, 5, 7 are all generators of this group.

4. $(\mathbb{Z}_n, \oplus)$ is a cyclic group of all $n \in \mathbb{N}$; 1 is a generator of this group.

In fact if $m \in \mathbb{Z}_n$ and $(m, n) = 1$ then m is a generator of this group.

5. $G = \{1, i, -1, -i\}$ is a cyclic group under usual multiplication; i is a generator, $-i$ is also a generator of G . However -1 is not a generator

Space for Hints

- of G since $\langle -1 \rangle = \{1, -1\} \neq G$.
6. $G = \{1, \omega, \omega^2\}$ where $\omega \neq 1$ is a cube root of unity is a cyclic group. ω and ω^2 are both generators of this group.
 7. In the group $G = (Z_7 - \{0\}, \odot)$, 3 and 5 are both generators. Here 2 is not a generator of G since $\langle 2 \rangle = \{2, 4, 1\} \neq G$.
 8. Let A be a set containing more than one element. Then $(\wp(A), \Delta)$ is not cyclic; for let $B \in \wp(A)$ be any element. Then $B \Delta B = \Phi$ so that $\langle B \rangle = \{B, \Phi\} \neq \wp(A)$.
 9. $(\mathbb{R}, +)$ is not a cyclic group since for any $x \in \mathbb{R}$, $\langle x \rangle = \{nx / n \in \mathbb{Z}\} \neq \mathbb{R}$.

Exercises:

Determine which of the following groups are cyclic. If it is cyclic find all the generators of the group.

1. $(6\mathbb{Z}, +)$.
2. $(\mathbb{Q}, +)$
3. The set of all n^{th} roots of unity under multiplication.
4. The group of symmetries of an equilateral triangle.
5. The group of symmetries of a rectangle.
6. The group of symmetries of a square.
7. $\{2^n / n \in \mathbb{Z}\}$ under usual multiplication.
8. $(Z_4, \oplus)$
9. $(\mathbb{R}^*, \cdot)$
10. $(Z_{11} - \{0\}, \odot)$
11. $G = \{e, p_1, p_2, p_3, p_4\}$ where

$$e = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}$$

$$p_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}$$

$$p_2 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 1 & 2 \end{pmatrix}$$

$$p_3 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 2 & 3 \end{pmatrix} \text{ and}$$

$$p_4 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix}$$

Answers:

1, 3, 7, 8, 10 and 11 are cyclic.

Theorem: 2.18

Any cyclic group is abelian.

Proof:

Let $G = \langle a \rangle$ be a cyclic group.

Let $x, y \in G$. Then $x = a^r$ and $y = a^s$ for some $r, s \in \mathbb{Z}$.

Hence $xy = a^r a^s = a^{r+s} = a^{s+r} = a^s a^r = yx$.

$\therefore G$ is abelian.

Theorem: 2.19

A subgroup of cyclic group is cyclic.

Proof:

Let G be a cyclic group generated by a and let H be a subgroup of G . We claim that H is cyclic.

Clearly every element of H is of the form a^n for some integer n .

Let m be the smallest positive integer such that $a^m \in H$. We claim that a^m is a generator of H .

Let $b \in H$. Then $b = a^n$ for some $n \in \mathbb{Z}$.

Let $n = mq + r$ where $0 \leq r < m$.

Then $b = a^n = a^{mq+r} = a^{mq} a^r = (a^m)^q a^r$.

$$\therefore a^r = (a^m)^{-q} b. \quad (1)$$

Now, $a^m \in H$. Since H is a subgroup, $(a^m)^{-q} \in H$.

Also $b \in H$.

By (1), $a^r \in H$ and $0 \leq r < m$.

But m is the least positive integer such that $a^n \in H$.

$\therefore r = 0$. Hence $b = a^n = a^{mq} = (a^m)^q$.

$\therefore$ Every element of H is a power of a^m .

$\therefore H = \langle a^m \rangle$ and hence H is cyclic.

Exercises:

1. Prove that if a is a generator of a cyclic group G then a^{-1} is also a generator of G .
2. Prove that any subgroup of $(\mathbb{Z}, +)$ is of the form $n\mathbb{Z}$ for some integer n .
3. Find the number of elements in the following cyclic subgroups.
 - (a) $\langle 2 \rangle$ in $(\mathbb{Z}_{18}, \oplus)$
 - (b) $\langle 18 \rangle$ in $(\mathbb{Z}_{30}, \oplus)$
 - (c) $\langle 5 \rangle$ in $(\mathbb{Z}_{80}, \oplus)$
 - (d) $\langle i \rangle$ in $\mathbb{C}^*$
4. Show that every proper subgroup of V_4 is cyclic. (However V_4 is not cyclic).
5. Show that every proper subgroup of S_3 is cyclic.
6. Give the multiplication table for the cyclic subgroup of S_5 generated by

$$p = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix}$$
7. Determine which of the following statements are true and which are false.
 - (a) For any fixed integer n , $\{kn + 1/k \in \mathbb{Z}\}$ is a subgroup of $(\mathbb{Z}, +)$.
 - (b) Every cyclic group is abelian.
 - (c) Every abelian group is cyclic.
 - (d) Every element of a cyclic group is a generator of the group.
 - (e) $(\mathbb{Z}, +)$ is a cyclic group.
 - (f) $(\mathbb{Q}, +)$ is a cyclic group.
 - (g) S_3 is a cyclic group.
 - (h) A_3 is a cyclic group.
 - (i) $(\mathbb{Z}_n, \oplus)$ is a cyclic group.
 - (j) $(\mathbb{Z}_n - \{0\}, \odot)$ is a cyclic group.

- (k) Any group of order 3 is cyclic.
 (l) Any group of order 4 is cyclic.
 (m) Given any positive integer n , there exists a cyclic group with n elements.
 (n) Every group has cyclic subgroups.
 (o) Every subgroup of a cyclic group is cyclic.
 (p) If every proper subgroup of a group G is cyclic then G is cyclic.
 (q) Every cyclic group has more than one generator.

Answers:

3. (a) 9 (b) 5 (c) 16 (d) 4

7. (b), (e), (h), (i), (k), (m), (n) and (o) are true.

Order of an Element:

1. Consider the group S_3 given in 3.4

$$p_1 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix};$$

$$p_1^2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = p_2.$$

$$p_1^3 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = e.$$

In this case 3 is the least positive integer such that $p_1^3 = e$

Also $\langle p_1 \rangle = \{e, p_1, p_2\}$ is a subgroup of S_3 of order 3.

2. Consider $(\mathbb{R}^+, \cdot)$. Form the sequence of elements $2, 2^2, 2^3, \dots, 2^n, \dots$. In this case there is no positive integer n such that $2^n = 1$ and $\langle 2 \rangle$ contains infinite numbers of elements.

Definition: 2.3.3

Let G be a group and let $a \in G$. The least positive integer n (if it exists) such that $a^n = e$ is called the **order** of a . If there is no positive

Space for Hints

integer n such that $a^n = e$, then the order of a is said to be infinite.

In example 1, p_1 is of order 3 and 2 is of infinite order in example 2.

In $(C^*, \cdot)$, i is an element of order 4.

Exercises:

1. Show that in any group G , e is the only element of order 1.
2. Find the order of -1 and 3 in $(\mathbb{Z}, +)$.
3. Find the order of -1 and 3 in $(\mathbb{R}^*, \cdot)$.
4. Find the order of -1 and $-i$ in $(C^*, \cdot)$.
5. Find the order of 2 and 3 in $(\mathbb{Z}_8, \oplus)$.
6. Show that in V_4 the order of every element other than the identity is 2.
7. Show that in $(\mathbb{Z}, +)$ the order of every element other than 0 is infinite.
8. Show that in $(\wp(S), \Delta)$ the order of every element other than Φ is 2.
9. Show that in $(C^*, \cdot)$ for every positive integer n there exists an element of order n .

Answers:

2. infinite
3. order of -1 is 2 and order of 3 is infinite.
4. order of -1 is 2 and order of $-i$ is 4.
5. order of 2 is 4 and order of 3 is 8.

Theorem: 2.20

Let G be a group and $a \in G$. Then the order of a is the same as the order of the cyclic group generated by a .

Proof:

Let a be an element of order n . Then $a^n = e$. We claim that $e, a, a^2, \dots, a^{n-1}$ are all distinct.

Suppose $a^r = a^s$ where $0 < r < s < n$.

Then $a^{s-r} = e$ and $s-r < n$ which contradicts the definition of the order of a . Hence $e, a, a^2, \dots, a^{n-1}$ are n distinct elements and $\langle a \rangle = \{e, a, a^2, \dots, a^{n-1}\}$ which is of order n .

If a is of infinite order, the sequence of elements $a, a^2, \dots, a^n, \dots$ are all distinct and are in $\langle a \rangle$. Hence $\langle a \rangle$ is an infinite group.

Theorem: 2.21

In a finite group every elements is an finite order.

Proof:

Let $a \in G$. If a is of infinite order, then $\langle a \rangle$ is an infinite subgroup of G , which is a contradiction since G is finite. Hence the order of a is finite.

Remark:

The converse of the above theorem is not true. (ie) if G is a group in which every element is of finite order then the group G need not be finite. For example, if S is any infinite set, then $(\wp(S), \Delta)$ is an infinite group. In this group $A \Delta A = \Phi$ for every $A \in \wp(S)$ so that the order of every element other than Φ is 2.

Theorem: 2.22

Let G be a group and a be an element of order n in G . Then $a^m = e$ iff n divides m .

Proof:

Suppose $n|m$. Then $m = nq$ where $q \in \mathbb{Z}$.

$$\therefore a^m = a^{nq} = (a^n)^q = e^q = e.$$

Conversely, let $a^m = e$.

Let $m = nq + r$ where $0 \leq r < n$.

$$\therefore a^m = a^{nq+r} = a^{nq} a^r = e a^r = a^r.$$

$$\therefore a^r = e \text{ and } 0 \leq r < n.$$

Now, since n is the smallest positive integer such that $a^n = e$, we have $r = 0$. Hence $m = nq$.

Therefore $n|m$.

Theorem: 2.23

Let G be a group and $a, b \in G$.

Then

(i) Order of a = order of a^{-1} .

(ii) Order of a = order of $b^{-1}ab$.

(iii) Order of ab = order of ba .

Space for Hints

Proof:

(i) Let a be an element of order n .

Then $a^n = e$.

$$\therefore (a^{-1})^n = (a^n)^{-1} = e^{-1} = e.$$

Now, if possible let $0 < m < n$ and $(a^{-1})^m = e$.

$\therefore (a^m)^{-1} = e$. Hence $a^m = e$ which contradicts the definition of the order of a . Thus n is the least positive integer such that $(a^{-1})^n = e$.

$\therefore$ The order of a^{-1} is n .

(ii) We shall first prove that for any positive integer r .

$$(b^{-1}ab)^r = b^{-1}a^r b. \quad (1)$$

(1) Is trivially true if $r = 1$.

Now, suppose that (1) is true for $r = k$ so that $(b^{-1}ab)^k = b^{-1}a^k b$.

Then

$$\begin{aligned} (b^{-1}ab)^{k+1} &= (b^{-1}ab)^k (b^{-1}ab) \\ &= (b^{-1}a^k b)(b^{-1}ab) \\ &= b^{-1}a^{k+1} b. \end{aligned}$$

Hence by induction (1) is true for all positive integers.

Now, let a be an element of order n . Then $a^n = e$.

$$\begin{aligned} \therefore (b^{-1}ab)^n &= b^{-1}a^n b \quad (\text{by (1)}) \\ &= b^{-1}eb = e. \end{aligned}$$

Now, if possible, let $0 < m < n$ and $(b^{-1}ab)^m = e$.

$\therefore b^{-1}a^m b = e$. Hence $a^m = e$ which contradicts the definition of the order of a . Thus n is the least positive integer such that $(b^{-1}ab)^n = e$.

$\therefore$ The order of $b^{-1}ab$ is n .

(iii) The order of ab = the order of $a^{-1}(ab)a$ (by (ii)) = the order of ba .

Theorem: 2.24

Let G be a group and let a be an element of order n in G . Then the order of a^s , where $0 < s < n$, is n/d where d is the g.c.d of n and s .

Proof:

Let $(n/d) = k$ and $(s/d) = \ell$ so that k and ℓ are relatively prime.

Now, $(a^s)^k = a^{sk} = a^{\ell dk} = a^{\ell n} = (a^n)^\ell = e$.

Further if m is any positive integer such that $(a^s)^m = e$ then $a^{sm} = e$.

Since order of a is n , we have $n | sm$.

$\therefore kd | \ell dm$. Hence $k | \ell m$

But k and ℓ are relatively prime.

Hence $k | m$ so that $m \geq k$.

Thus k is the least positive integer such that $(a^s)^k = e$.

$\therefore$ order of $a^s = k = n/d$.

Corollary: 1

The order of any power of a cannot exceed the order of a .

Corollary: 2

Let G be a finite cyclic group of order n generated by an element a . Then a^s generates a cyclic group of order n/d where d is the g.c.d of n and s .

Corollary: 3

Let G be a finite cyclic group of order n generated by an element a . a^s is a generator of G iff s and n are relatively prime. Hence the number of generators of a cyclic group of order n is $\phi(n)$ where $\phi(n)$ is the number of positive integers less than n relatively prime to n .

For example, consider the group $(\mathbb{Z}_{12}, \oplus)$.

$\phi(12) = 4$. Hence the group has exactly 4 generators and they are 1, 5, 7 and 11.

Problems:

Problem: 1

If G is a finite group with even number of elements then G contains at least one element of order 2.

Solution:

a is an element of order 2 $\Leftrightarrow a^2 = e$

$\Leftrightarrow a^{-1} = a$.

Space for Hints

Hence it is enough if we prove that there exists an element different from e in G whose inverse is itself.

Let $S = \{a / a \in G, a \neq a^{-1}\}$.

Clearly $a \in S \Rightarrow a^{-1} \in S$ and $a \neq a^{-1}$.

Hence S contains an even number of elements.

Also $e \notin S$.

Hence $S \cup \{e\}$ contains an odd number of elements. Since the order of the group is even, there exists at least one element $a \notin S \cup \{e\}$. Clearly $a = a^{-1}$.

Problem: 2

The order of a permutation p is the $\ell.c.m$ of the lengths of its disjoint cycles.

Solution:

Let $p = c_1 c_2 \dots c_r$ where the c_i 's are mutually disjoint cycles of lengths ℓ_i . Now, let $p^m = e$.

Since product of disjoint cycles is commutative,

$$e = p^m = (c_1 c_2 \dots c_r)^m = c_1^m c_2^m \dots c_r^m.$$

Now, since the elements moved by one cycle are left fixed by all the other cycles, $c_1^m = c_2^m = \dots c_r^m = e$.

Now, $c_1^m = e \Rightarrow \ell_1 \mid m$ since the order of $c_1 = \ell_1$.

Similarly $\ell_2, \ell_3, \dots, \ell_r$ divide m .

Thus m is a common multiple of $\ell_1, \ell_2, \dots, \ell_r$.

$\therefore$ The order of p is the least such m which is obviously the $\ell.c.m.$ of $\ell_1, \ell_2, \dots, \ell_r$.

Problem: 3

If a is a generator of the cyclic group G and if there exist two unequal integers m and n such that $a^m = a^n$, prove that G is a finite group.

Solution:

Since m and n are unequal we may assume that $m > n$.

Hence $m - n$ is a positive integer.

$$\text{Also } a^m = a^n \Rightarrow a^{m-n} = e.$$

$\therefore$ Order of a is finite.

$\therefore G = \langle a \rangle$ is a finite group (by theorem 2.20)

Exercises:

1. Show that a group G of order n is cyclic iff G contains an element of order n .
2. Find the number of generators of the cyclic groups of order 8, 24 and 60.
3. Let p and q be prime number. Find the number of generators of Z_{pq} .
4. Find the number of generators of Z_p , where p is prime.
5. Find two elements a, b in a group such that
 - (a) Order of $ab \neq (\text{order of } a)(\text{order of } b)$.
 - (b) Order of $ab = (\text{order of } a)(\text{order of } b)$.
6. Find the order of the following permutations.
 - (a) $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 1 \end{pmatrix}$
 - (b) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 7 & 6 & 4 & 1 & 2 & 3 \end{pmatrix}$
 - (c) $(12345)(67)(1657)$
 - (d) $(12)(23)(345)(1456)$.
7. Find all the elements of finite order in the following groups.
 - (a) $(\mathbb{Z}, +)$
 - (b) $(\mathbb{R}^*, \cdot)$
 - (c) $(\wp(S), \Delta)$
8. Determine which of the following are true and which are false.
 - (a) In a finite group the order of every element is finite.
 - (b) If in a group every element is of finite order then the group is finite.
 - (c) In an infinite group the order of every element is infinite.
 - (d) The order of e is zero.
 - (e) The order of every element of $(\mathbb{R}^*, \cdot)$ is finite.

Answers:

2. 4, 8 and 16.

6. (a) 4 (b) 4 (c) 7 (d) 6

7. (a) 0 (b) 1 and -1 (c) All elements.

8. (a) T (b) F (c) F (d) F (e) F

COSETS AND LAGRANGE'S THEOREM

3.1 Introduction:

Consider S_3 . Let $H = \{e, p_3\}$. H is a subgroup of S_3 . This subgroup does not contain the elements p_1, p_2, p_4 and p_5 . Let us now perform the binary operation between p_1 and each element of H . We denote the resultant set by the symbol p_1H .

$$\text{Thus } p_1H = \{p_1e, p_1p_3\} = \{p_1, p_4\}.$$

Now, the element p_2 belongs neither to H nor to p_1H . Therefore, we now perform the binary operation between p_2 and the element of H .

Thus $p_2H = \{p_2e, p_2p_3\} = \{p_2, p_5\}$. The union of the three sets H, p_1H, p_2H gives all the elements of S_3 (ie) $S_3 = H \cup p_1H \cup p_2H$. Further H, p_1H and p_2H are mutually disjoint. Hence $\{H, p_1H, p_2H\}$ is a partition of S_3 .

Definition: 3.1.1

Let H be a subgroup of a group G . Let $a \in G$. Then the set $aH = \{ah / h \in H\}$ is called the **left coset** of H defined by a in G .

Similarly $Ha = \{ha / h \in H\}$ is called the **right coset** of H defined by a .

Examples:

1. Let us determine the left cosets of $(5Z, +)$ in $(Z, +)$. Here the operation is $+$. $0 + 5Z = 5Z$ is itself a left coset.

Another left coset is $1 + 5Z = \{1 + 5n / n \in Z\}$. We notice that this left coset contains all integers having remainder 1 when divided by 5.

Similarly

$$2 + 5Z = \{2 + 5n / n \in Z\}$$

$$3 + 5Z = \{3 + 5n / n \in Z\}.$$

$$\text{and } 4 + 5Z = \{4 + 5n / n \in Z\}.$$

These are all the left cosets of $(5Z, +)$. Here also we note that all the left cosets are mutually disjoint, and their union is Z . In other words the collection of all left cosets forms a partition of the group.

2. Consider $(\mathbb{Z}_{12}, \oplus)$. Then $H = \{0, 4, 8\}$ is a subgroup of G .

The left cosets of H are given by

$$0 + H = \{0, 4, 8\} = H$$

$$1 + H = \{1, 5, 9\}$$

$$2 + H = \{2, 6, 10\}$$

and $3 + H = \{3, 7, 11\}$

we notice that

$$4 + H = \{4, 8, 0\} = H \text{ and}$$

$$5 + H = \{5, 9, 1\} = 1 + H \text{ etc.}$$

Exercises:

1. Find all the left cosets of $(n\mathbb{Z}, +)$ in $(\mathbb{Z}, +)$.
2. Find all the left cosets of $\{0, 3, 6, 9\}$ in $(\mathbb{Z}_{12}, \oplus)$.
3. Find all the left cosets of $\{1, 6\}$ in $(\mathbb{Z}_7 - \{0\}, \odot)$.

Theorem: 3.1

Let G be a group and H be a subgroup of G . Then

- (i) $a \in H \Rightarrow aH = H$
- (ii) $aH = bH \Rightarrow a^{-1}b \in H$.
- (iii) $a \in bH \Rightarrow a^{-1} \in Hb^{-1}$.
- (iv) $a \in bH \Rightarrow aH = bH$.

Proof:

- (i) Let $a \in H$. We claim that $aH = H$. Let $x \in aH$. Then $x = ah$ for some $h \in H$. Now, $a \in H$ and $h \in H \Rightarrow ah = x \in H$ (since H is a subgroup).

Hence $aH \subseteq H$.

Let $x \in H$. Then $x = a(a^{-1}x) \in aH$.

Hence $H \subseteq aH$. Thus $H = aH$.

Conversely, let $aH = H$. Now $a = ae \in aH$.

$\therefore a \in H$.

- (ii) Let $aH = bH$.

$$\therefore a^{-1}(aH) = a^{-1}(bH).$$

$$\therefore H = (a^{-1}b)H.$$

$$\therefore a^{-1}b \in H \text{ (by i).}$$

Conversely let $a^{-1}b \in H$.

Then $a^{-1}bH = H$ (by i).

$$\therefore aa^{-1}bH = aH \text{ and hence } bH = aH.$$

(iii) Let $a \in bH$. Then $a = bh$ for some $h \in H$.

$$\therefore a^{-1} = (bh)^{-1} = h^{-1}b^{-1} \in Hb^{-1}.$$

Converse can be similarly proved.

(iv) Let $a \in bH$. We claim that $aH = bH$.

Let $x \in aH$. Then $x = ah_1$ for some $h_1 \in H$. Also

$$a \in bH \Rightarrow a = bh_2 \text{ for some } h_2 \in H \dots (1)$$

$$\therefore x = ((bh_2)h_1) = b(h_2h_1) \in bH.$$

$$\therefore aH \subseteq bH.$$

Now, let $x \in bH$. Then $x = bh_3$ for some $h_3 \in H$.

Also from (1), $b = ah_2^{-1}$.

$$\therefore x = ah_2^{-1}h_3 \in aH.$$

$$\therefore bH \subseteq aH. \text{ Hence } aH = bH.$$

Conversely, let $aH = bH$.

Then $a = ae \in aH$.

$$\therefore a \in bH.$$

Theorem: 3.2

Let H be a subgroup of G . Then

(i) Any two left cosets of H are either identical or disjoint.

(ii) Union of all the left cosets of H is G .

(iii) The number of elements in any left coset aH is the same as the number of elements in H .

Proof:

(i) Let aH and bH be two left cosets.

Suppose aH and bH are not disjoint. We claim that $aH = bH$.

Since aH and bH are not disjoint, $aH \cap bH \neq \Phi$.

$$\therefore \text{There exists an element } c \in aH \cap bH.$$

$$\therefore c \in aH \text{ and } c \in bH.$$

$\therefore aH = cH$ and $bH = cH$ (by (iv) of Theorem 3.1)

$\therefore aH = bH$.

(ii) Let $a \in G$. Then $a = ae \in aH$.

$\therefore$ Every element of G belongs to a left coset of H .

$\therefore$ The union of all the left cosets of H is G .

(iii) The map $f: H \rightarrow aH$ defined by $f(h) = ah$ is clearly a bijection. Hence every left coset has the same number of elements as H .

Note: 1

This theorem shows that the collection of all left cosets forms a partition of the group.

Note: 2

The above result is true if we replace left cosets by right cosets. In what follows, the results we prove for left cosets are also true for right cosets.

Remark:

Let H be a subgroup of G . We define a relation in G as follows.

Define $a \sim b \Leftrightarrow a^{-1}b \in H$.

Then $\sim$ is an equivalence relation.

For, $a^{-1}a = e \in H$. Hence $a \sim a$.

Hence $\sim$ is reflexive.

$$a \sim b \Rightarrow a^{-1}b \in H \Rightarrow (a^{-1}b)^{-1} \in H.$$

$$\Rightarrow b^{-1}a \in H \Rightarrow b \sim a.$$

$$\therefore a \sim b \Rightarrow b \sim a.$$

Hence $\sim$ is symmetric.

Now,

$$a \sim b \text{ and } b \sim c \Rightarrow a^{-1}b \in H \text{ and } b^{-1}c \in H$$

$$\Rightarrow (a^{-1}b)(b^{-1}c) \in H$$

$$\Rightarrow a^{-1}c \in H$$

$$\Rightarrow a \sim c.$$

Hence $\sim$ is transitive.

Thus $\sim$ is an equivalence relation.

Now, we claim that equivalence class $[a] = aH$.

Let $b \in [a]$. Then $b \sim a$.

$$\therefore a^{-1}b \in H.$$

$$\therefore a^{-1}b = h \text{ for some } h \in H.$$

$$\therefore b = ah. \text{ Hence } b \in aH.$$

$$\therefore [a] \subseteq aH.$$

Also,

$$b \in aH \Rightarrow b = ah \text{ for some } h \in H.$$

$$\Rightarrow a^{-1}b = h \in H.$$

$$\Rightarrow a \sim b$$

$$\Rightarrow b \in [a].$$

Thus the left cosets of H in G are precisely the equivalence classes determined by $\sim$. Hence the left cosets form a partition of G . This gives another proof of Theorem 3.2.

Theorem: 3.3

Let H be a subgroup of G . The number of left cosets of H is the same as the number of right cosets of H .

Proof:

Let L and R respectively denote the set of left and right cosets of H .

We define a map $f : L \rightarrow R$ by $f(aH) = Ha^{-1}$.

f is well defined. For $aH = bH \Rightarrow a^{-1}b \in H$

$$\Rightarrow a^{-1} \in Hb^{-1}$$

$$\Rightarrow Ha^{-1} = Hb^{-1}$$

f is 1-1. For,

$$f(aH) = f(bH)$$

$$\Rightarrow Ha^{-1} = Hb^{-1}$$

$$\Rightarrow a^{-1} \in Hb^{-1}$$

$$\Rightarrow a^{-1} = hb^{-1} \text{ for some } h \in H.$$

$$\Rightarrow a = bh^{-1}$$

$$\Rightarrow a \in bH$$

$$\Rightarrow aH = bH.$$

f is **onto**. For, every right coset Ha has a pre – image under f namely $a^{-1}H$.

Hence f is a bijection from L to R . Hence the number of left cosets is the same as the number of right cosets.

Definition: 3.1.2

Let H be a subgroup of G . The number of distinct left (right) cosets of H in G is called the **index** of H in G is denoted by $[G: H]$.

Example:

In $(\mathbb{Z}_8, \oplus)$, $H = \{0, 4\}$ is a subgroup. The left cosets of H are given by

$$0 + H = \{0, 4\} = H$$

$$1 + H = \{1, 5\}.$$

$$2 + H = \{2, 6\}.$$

$$3 + H = \{3, 7\}.$$

These are the four distinct left cosets of H .

Hence the index of the subgroup H is 4.

Note that $[\mathbb{Z}_8 : H] \times |H| = 4 \times 2 = 8 = |\mathbb{Z}_8|$.

Exercises:

1. Find the index of $(n\mathbb{Z}, +)$ in $(\mathbb{Z}, +)$.
2. Find the index of $(8\mathbb{Z}, +)$ in $(2\mathbb{Z}, +)$.
3. Find the index of $\{0, 3, 6, 9\}$ in $(\mathbb{Z}_{12}, \oplus)$.
4. Find the index of $\{1, 6\}$ in $(\mathbb{Z}_7 - \{0\}, \odot)$
5. Determine which of the following statements are true and which are false.

Let G be a group and H a subgroup of G , Then

- (a) H itself is a left coset of H .
- (b) The identity element belongs to every left coset of H .
- (c) Every element of H belongs to at least one left coset of H .
- (d) Every element of H belongs to exactly one left coset of H .
- (e) The number of left cosets of H is the same as the number of right coset of H .
- (f) If $b \in Ha$ then $Ha = Hb$.

Space for Hints

(g) If $c \in Ha \cap Hb$ then $Ha = Hb = Hc$.

(h) $a \in H$ if and only if $aH = H$.

Answers:

- | | | | | |
|----|-------|-------|-------|-------|
| 5. | (a) T | (b) F | (c) T | (d) T |
| | (e) T | (f) T | (g) T | (h) T |

Theorem: 3.4 (Lagrange's theorem)

Let G be a finite group of order n and H be any subgroup of G . Then the order of H divides the order of G .

Proof:

Let $|H| = m$ and $[G : H] = r$.

Then the number of distinct left cosets of H in G is r .

By theorem 3.2, these r left cosets are mutually disjoint, they have the same number of elements namely m and their union is G .

$\therefore n = rm$. Hence m divides n .

Corollary:

$$[G : H] = \frac{|G|}{|H|}$$

Note: 1

Lagrange's theorem has many important applications in group theory. For example, a group G of order 8 cannot have subgroups of order 3, 5, 6 or 7. In fact any proper subgroup of G must be of order 2 or 4.

Note: 2

Any group of prime order has no proper subgroups.

Note: 3

The converse of Lagrange's theorem is false. (i.e) If G is a group of order n and m divides n , then G need not have a subgroup of order m .

For example A_4 is a group of order 12 and it does not have a subgroup of order 6.

However there are groups in which the converse of Lagrange's theorem is true.

For example, consider S_3 . This is a group of order 6. $\{e, p_4\}$ is a subgroup of order 2 and $\{3, p_1, p_2\}$ is a subgroup of order 3. Hence for every

divisor m of 6, there is a subgroup of S_3 of order m .

Exercise:

1. Can a group of order 12 contain a subgroup of order 8?
2. Show that the converse of Lagrange's theorem is true in V_4 .
3. Show that the converse of Lagrange's theorem is true in any finite cyclic group.

Theorem: 3.5

The order of any element of a finite group G divides the order of G .

Proof:

Let G be a group of order n . Let $a \in G$ be an element of order m . Then the order of a is the same as the order m . Then the order of a is the same as the order of the cyclic group $\langle a \rangle$.

Now, by Lagrange's theorem the order of the subgroup $\langle a \rangle$ divides the order of G . Hence $m|n$.

Theorem: 3.6

Every group of prime order is cyclic.

Proof:

Let G be a group of order p where p is prime. Let $a \in G$ and $a \neq e$.

By Theorem 3.5 order of a divides p .

$\therefore$ Order of a is 1 or p .

Since $a \neq e$ order of a is p .

Hence $G = \langle a \rangle$ so that G is cyclic.

Theorem: 3.7

Let G be a group of order n . Let $a \in G$ then $a^n = e$.

Proof:

Let the order of a be m . Then m divides n . Hence $n = mq$.

$\therefore a^n = a^{mq} = (a^m)^q = e^q = e$.

Theorem: 3.8 (Euler's theorem)

If n is any integer and $(a, n) = 1$ then $a^{\phi(n)} \equiv 1 \pmod{n}$.

($\phi(n)$ is the number of positive integers less than n relatively prime to n)

Proof:

Let $G = \{m / m < n \text{ and } (m, n) = 1\}$. G is a group under multiplication modulo n . This group is of order $\phi(n)$.

Now, let $(a, n) = 1$.

Let $a = qn + r; 0 \leq r < n$ so that $a \equiv r \pmod{n}$.

Since $(a, n) = 1$ we have $(n, r) = 1$ so that $r \in G$.

$$\therefore r^{\phi(n)} = 1 \text{ (by theorem 3.7)}$$

$$\therefore r^{\phi(n)} \equiv 1 \pmod{n}.$$

Also $a^{\phi(n)} \equiv r^{\phi(n)} \pmod{n}$ so that $a^{\phi(n)} \equiv 1 \pmod{n}$ (since ' $\equiv$ ' is transitive).

Theorem: 3.9 (Fermat's theorem)

Let p be a prime number and a be any integer relatively prime to p . Then $a^{p-1} \equiv 1 \pmod{p}$.

Proof:

Since p is prime, $\phi(p) = p - 1$ and hence the result follows from Euler's theorem.

Theorem: 3.10

A group G has no proper subgroups if it is a cyclic group of prime order.

Proof:

Suppose G is a group of prime order. Then it follows from Lagrange's theorem that G has no proper subgroups.

Conversely, let G be a group having no proper subgroup. First we shall prove that G is cyclic.

Suppose G is not cyclic. Let $a \in G$ and $a \neq e$.

Then the cyclic group $\langle a \rangle$ is a proper subgroup of G which is a contradiction. Hence G is cyclic.

Also G cannot be infinite, for, an infinite cyclic group contains a proper subgroup $\langle a^2 \rangle$. Hence G must be of finite order, say, n .

We claim that n is prime. If possible let n be a composite number. Let $n = pq$ where $p, q > 1$.

Let $a \in G$ be a generator of the group.

Then $\langle a^p \rangle$ is a subgroup of order q and hence is a proper subgroup of G which is a contradiction.

Hence n is prime.

$\therefore G$ is a cyclic group of prime order.

Problems:

problem: 1

Let A and B be subgroups of a finite group G such that A is a subgroup of B . Show that

$$[G : A] = [G : B][B : A].$$

Solution:

$$[G : A] = \frac{|G|}{|A|} \text{ (by Lagrange's theorem)}$$

$$[G : B] = \frac{|G|}{|B|}$$

$$\text{and } [B : A] = \frac{|B|}{|A|}$$

$$\therefore [G : B][B : A] = \frac{|G| \cdot |B|}{|B| \cdot |A|} = \frac{|G|}{|A|} = [G : A].$$

Problem: 2

Let A and B be two finite subgroups of a group G such that $|A|$ and $|B|$ have no common divisors. Then show that $A \cap B = \{e\}$.

Solution:

$A \cap B$ is a subgroup of A and B .

$\therefore$ By Lagrange's theorem, $|A \cap B|$ divides $|A|$ and $|B|$. But by hypothesis $|A|$ and $|B|$ have no common divisors.

$$\therefore |A \cap B| = 1. \text{ Hence } A \cap B = \{e\}.$$

Problem: 3

Let H and K be two subgroups of G of finite index in G . Prove that $H \cap K$ is a subgroup of finite index in G .

Solution:

By theorem 2.10 $H \cap K$ is a subgroup of finite index in G .

Let $[G : H] = m$ and $[G : K] = n$.

We claim that for any $a \in G$, $(H \cap K)a = Ha \cap Ka$.

Clearly, $H \cap K \subseteq H$ and K .

$$\therefore (H \cap K)a \subseteq Ha \text{ and } Ka.$$

$$\therefore (H \cap K)a \subseteq Ha \cap Ka. \quad (1)$$

Now, let $x \in Ha \cap Ka$.

$$\therefore x \in Ha \text{ and } x \in Ka$$

$$\therefore x = ha \text{ for some } h \in H \text{ and}$$

$$x = ka \text{ for some } k \in K.$$

$$\therefore x = ha = ka$$

$$\therefore h = k$$

$$\therefore h \in H \cap K.$$

$$\therefore x \in (H \cap K)a.$$

$$\therefore Ha \cap Ka \subseteq (H \cap K)a. \quad (2)$$

From (1) and (2) we have

$$(H \cap K)a = Ha \cap Ka.$$

$\therefore$ Every right coset of $H \cap K$ in G is the intersection of a right coset of H and a right coset of K .

Also since $[G : H] = m$, the number of right cosets of H in G is m .

Similarly the number of right cosets of K in G is n .

Hence the number of right cosets of $H \cap K$ in G is at most mn .

$$\therefore [G : H \cap K] \leq mn.$$

$\therefore H \cap K$ is a subgroup of finite index in G .

Problem: 4

Let H and K be two finite subgroups of a group G . Then

$$|HK| = \frac{|H| |K|}{|H \cap K|}.$$

Proof:

Let $L = H \cap K$. Since H and K are subgroups of G , L is also a subgroup of G and $L \subseteq H$ and K .

Now, let $Lx_1, Lx_2, \dots, Lx_m$ be the distinct right cosets of L in K so that

$$K = Lx_1 \cup Lx_2 \cup \dots \cup Lx_m \quad (1)$$

and

$$m = [K : L] = \frac{|K|}{|L|} = \frac{|K|}{|H \cap K|} \quad (2)$$

From (1)

$$\begin{aligned} HK &= HLx_1 \cup HLx_2 \cup \dots \cup HLx_m \\ &= Hx_1 \cup Hx_2 \cup \dots \cup Hx_m \text{ (since } L \subseteq H \text{)} \end{aligned} \quad (3)$$

We claim that the cosets $Hx_1, Hx_2, \dots, Hx_m$ are distinct.

Suppose $Hx_i = Hx_j$.

$$\therefore x_i x_j^{-1} \in H$$

Also $x_i, x_j \in K$ and hence $x_i x_j^{-1} \in K$.

$$\therefore x_i x_j^{-1} \in H \cap K = L.$$

Hence $Lx_i = Lx_j$ which is a contradiction since the cosets $Lx_1, Lx_2, \dots, Lx_m$ are distinct.

Thus from (3) we have

$$\begin{aligned} |HK| &= |Hx_1| + |Hx_2| + \dots + |Hx_m| \\ &= m|H| \\ &= \frac{|H||K|}{|H \cap K|} \text{ by (2).} \end{aligned}$$

Problem: 5

Let H and K be two subgroups of a finite group G such that $|H| > \sqrt{|G|}$ and $|K| > \sqrt{|G|}$. Then $H \cap K \neq \{e\}$.

Proof:

Suppose $H \cap K = \{e\}$.

$$\therefore |H \cap K| = 1$$

$$\begin{aligned} \therefore |HK| &= \frac{|H||K|}{|H \cap K|} \text{ (by Problem 4)} \\ &= |H||K| \end{aligned}$$

$$> \sqrt{|G|} \sqrt{|G|} = |G|.$$

$\therefore |HK| > |G|$ which is a contradiction.

$$\therefore H \cap K \neq \{e\}.$$

Exercises:

1. Show that any group of order 17 is cyclic.
2. Show that any infinite group has proper subgroups.
3. Prove that in an infinite cyclic group all subgroups other than $\{e\}$ are infinite.
4. Is it possible for an infinite group to have a finite subgroup?
5. Let G be a finite group and H a subgroup of G . What is H if
 - (i) $[G : H] = 1$
 - (ii) $[G : H] = |G|$.
6. Determine which of the following statements are true and which are false.
 - (a) The index of any subgroup of an infinite group is infinite.
 - (b) Any group having no proper subgroup is cyclic.
 - (c) Every group of prime order is abelian.
 - (d) Every group of prime order is cyclic.
 - (e) A subgroup of a group is a left coset of itself.
 - (f) In a finite group the order of any element divides the order of the group.
 - (g) For any subgroup of a group, the number of left cosets is the same as the number of right cosets.
 - (h) Z_{12} contains a subgroup of order 3.
 - (i) Any infinite group has proper subgroups.
 - (j) Any group which is not cyclic has proper subgroups.

Answers:

5. (a) $H = G$. (b) $H = \{e\}$
6. (a) is false and the other statements are true.

3.2 Normal Subgroups and Quotient Groups:

Consider the subgroup $H = \{e, p_3\}$ of S_3 . Then $Hp_1 = \{p_1, p_5\}$ and $p_1H = \{p_1, p_4\}$. Hence $Hp_1 \neq p_1H$. Thus a left coset need not be equal to the

corresponding right coset. However there are subgroups for which every left coset is the same as the corresponding right coset.

For example, consider the subgroup.

$$H = \{e, p_1, p_2\} \text{ of } S_3.$$

Clearly $Ha = aH = H = \{e, p_1, p_2\}$ for all $a \in H$ and

$$aH = Ha = S_3 - H = \{p_3, p_4, p_5\} \text{ for all } a \notin H.$$

$$\therefore aH = Ha \text{ for all } a \in S_3.$$

Thus for some subgroups every left coset is a right coset. This leads to the following definition of a special class of subgroups.

Definition: 3.2.1

A subgroup H of G is called a **normal subgroup** of G if $aH = Ha$ all $a \in G$.

Examples:

1. For any group G , $\{e\}$ and G are normal subgroups.
2. In S_3 , the subgroup $\{e, p_1, p_2\}$ is normal.
3. In S_3 , the subgroup $\{e, p_3\}$ is not a normal subgroup.

Theorem: 3.11

Every subgroup of an abelian group is a normal subgroup.

Proof:

Let G be an abelian group and let H be a subgroup of G . Let $a \in G$.

We claim that $aH = Ha$.

Let $x \in aH$. Then

$$\begin{aligned} x &= ah \text{ for some } h \in H. \\ &= ha \text{ (since } G \text{ is abelian).} \end{aligned}$$

$$\therefore x \in Ha. \text{ Hence } aH \subseteq Ha.$$

Similarly $Ha \subseteq aH$.

$$\therefore aH = Ha \text{ and hence } H \text{ is a normal subgroup of } G.$$

Examples:

- (i). $n\mathbb{Z}$ is a normal subgroup of $(\mathbb{Z}, +)$.
- (ii). Every subgroup of $(\mathbb{Z}_n, \oplus)$ is normal.

- (iii). Since any cyclic group is abelian any subgroup of a cyclic group is normal.

Theorem: 3.12

Let H be a subgroup of index 2 in a group G . Then H is a normal subgroup of G .

Proof:

If $a \in H$ then $H = aH = Ha$.

If $a \notin H$, then aH is a left coset different from H .

Hence $H \cap aH = \Phi$.

Further, since index of H in G is 2, $H \cup aH = G$.

Hence $aH = G - H$.

Similarly $Ha = G - H$ so that $aH = Ha$.

Hence H is a normal subgroup of G .

Example:

The alternating group A_n is a subgroup of index 2 in S_n and hence is a normal subgroup of S_n .

Theorem: 3.13

Let N be a subgroup of G . Then the following are equivalent.

- (i). N is a normal subgroup of G .
- (ii). $aNa^{-1} = N$ for all $a \in G$.
- (iii). $aNa^{-1} \subseteq N$ for all $a \in G$.
- (iv). $ana^{-1} \in N$ for all $n \in N$ and $a \in G$.

Proof:

(i) $\Rightarrow$ (ii)

Suppose N is a normal subgroup of G .

$\therefore aN = Na$ for all $a \in G$.

$\therefore aNa^{-1} = Naa^{-1} = Ne = N$.

(ii) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (iv) are obvious.

(iv) $\Rightarrow$ (i).

Suppose that $ana^{-1} \in N$ for all $n \in N$ and $a \in G$. We claim that $aN = Na$.

Let $x \in aN$.

$$\therefore x = an \text{ for some } n \in N.$$

$$\therefore x = (ana^{-1})a \in Na \text{ (since } ana^{-1} \in N).$$

$$\therefore aN \subseteq Na \quad (1)$$

Now, let $x \in Na$.

$$\therefore x = na \text{ for some } n \in N.$$

$$\therefore x = a(a^{-1}na) = a(a^{-1}n(a^{-1})^{-1}) \in aN.$$

$$\therefore Na \subseteq aN \quad (2)$$

From (1) and (2) we get $Na = aN$.

Hence N is a normal subgroup of G .

Problems:

Problem: 1

Prove that the intersection of two normal subgroups of a group G is a normal subgroup of G .

Solution:

Let H and K be two normal subgroups of G . Then $H \cap K$ is a subgroup of G .

Now, let $a \in G$ and $x \in H \cap K$. Then $x \in H$ and $x \in K$.

Since H and K are normal $axa^{-1} \in H$ and $axa^{-1} \in K$.

Hence $axa^{-1} \in H \cap K$. Thus $H \cap K$ is a normal subgroup of G .

Problem: 2

The center H of a group G is a normal subgroup of G .

Solution:

The center H of G is given by

$$H = \{a / a \in G, ax = xa \text{ for all } x \in G\}.$$

Now let $x \in H$ and $a \in G$. Hence $ax = xa$.

$$\therefore x = axa^{-1} \in H.$$

Hence H is a normal subgroup of G .

Problem: 3

Let H be a subgroup of G . Let $a \in G$. Then aHa^{-1} is a subgroup of G .

Solution:

$$e = aea^{-1} \in aHa^{-1} \text{ and hence } aHa^{-1} \neq \Phi.$$

Now, let $x, y \in aHa^{-1}$.

Then $x = ah_1a^{-1}$ and $y = ah_2a^{-1}$ where $h_1, h_2 \in H$.

Now,

$$\begin{aligned} xy^{-1} &= (ah_1a^{-1})(ah_2a^{-1})^{-1} \\ &= (ah_1a^{-1})(ah_2^{-1}a^{-1}) \\ &= a(h_1h_2^{-1})a^{-1} \in aHa^{-1} \end{aligned}$$

$\therefore aHa^{-1}$ is a subgroup of G .

Problem: 4

Show that if a group G has exactly one subgroup H of given order, then H is a normal subgroup of G .

Solution:

Let the order of H be m .

Let $a \in G$. Then by solved problem 3, aHa^{-1} is also a subgroup of G .

We claim that $|H| = |aHa^{-1}| = m$.

Now, consider $f: H \rightarrow aHa^{-1}$ defined by $f(h) = aha^{-1}$.

f is 1-1, for,

$$\begin{aligned} f(h_1) = f(h_2) &\Rightarrow ah_1a^{-1} = ah_2a^{-1} \\ &\Rightarrow h_1 = h_2. \end{aligned}$$

f is onto, for, let $x = aha^{-1} \in aHa^{-1}$.

Then $f(h) = x$. Thus f is a bijection.

$$\therefore |H| = |aHa^{-1}| = m.$$

But H is the only subgroup of G of order m .

$$\therefore aHa^{-1} = H. \text{ Hence } aH = Ha.$$

$$\therefore H \text{ is a normal subgroup of } G.$$

Problem: 5

Show that if H and N are subgroup of a group G and N is normal in G , then $H \cap N$ is normal in H . Show by an example that $H \cap N$ need not be normal in G .

Solution:

Let $x \in H \cap N$ and $a \in H$.

We claim that $axa^{-1} \in H \cap N$.

Now, $x \in N$ and $a \in H \Rightarrow axa^{-1} \in N$ (since N is a normal subgroup).

Also $x \in N$ and $a \in H \Rightarrow axa^{-1} \in H$ (since H is a group).

Hence $axa^{-1} \in H \cap N$.

$\therefore H \cap N$ is a normal subgroup of H .

The following example shows that $H \cap N$ need not be normal in G .

Let $G = S_3$. Take $N = G$ and $H = \{e, p_3\}$.

Now $H \cap N = H$ which is not normal in G .

Problem: 6

If H is a subgroup of G and N is a normal subgroup of G then

HN is a subgroup of G .

Solution:

To prove that HN is a subgroup of G , it is enough if we prove that $HN = NH$.

Let $x \in HN$. Then $x = hn$ where $h \in H$ and $n \in N$.

$\therefore x \in hN$.

But $hN = Nh$ (since N is normal).

$\therefore x \in Nh$. Hence $x = n_1h$ where $n_1 \in N$.

$\therefore x \in NH$. Hence $HN \subseteq NH$.

Similarly $NH \subseteq HN$.

$\therefore HN = NH$. Hence HN is a subgroup of G .

Problem: 7

M and N are normal subgroups of a group G such that $M \cap N = \{e\}$. Show that every element of M commutes with every element of N .

Solution:

Let $a \in M$ and $b \in N$.

We claim that $ab = ba$.

Consider the element $aba^{-1}b^{-1}$.

Since $a^{-1} \in M$ and M is normal, $ba^{-1}b^{-1} \in M$.

Also $a \in M$, so that $aba^{-1}b^{-1} \in M$.

Again, since $b \in N$ and N is normal, $aba^{-1} \in N$.

Also, $b^{-1} \in N$, so that $aba^{-1}b^{-1} \in N$.

Thus $aba^{-1}b^{-1} \in M \cap N = \{e\}$.

$\therefore aba^{-1}b^{-1} = e$, so that $ab = ba$.

Exercises:

1. If A is normal in G and B a subgroup of G such that $A \subseteq B \subseteq G$, then prove that A is a normal subgroup of B .
2. Show that a subgroup of index 3 need not be normal.
3. Let A, B, C, D be subgroups of G . Let A be a normal subgroup of B , and C a normal subgroup of D . Then show that $A \cap C$ is normal subgroup of $B \cap D$.
4. Let G, G' be two groups with identity e and e' respectively. Prove that the subsets $G \times \{e'\}$ and $\{e\} \times G'$ are normal subgroups of $G \times G'$.
5. Let H and N be normal subgroups of a group G . Prove that HN is a normal subgroups of G .

Let G be a group and N be a subgroup of G . We denote by G/N the set of all right cosets of N in G . Thus $G/N = \{Na / a \in G\}$. The following theorem enables us to introduce a group structure on G/N when N is a normal subgroup of G .

Theorem: 3.14

A subgroup N of G is normal iff the product of two right cosets of N is again a right coset of N .

Proof:

Suppose N is a normal subgroup of G .

Then

$$\begin{aligned} NaNb &= N(aN)b \\ &= N(Nab) \text{ (since } aN = Na \text{)} \\ &= NNab \\ &= Nab \text{ (since } NN = N \text{)}. \end{aligned}$$

Conversely suppose that the product of any two right cosets of N is again a right coset of N . Then $NaNb$ is a right coset of N .

Further $ab = (ea)(eb) \in NaNb$.

Hence $NaNb = Nab$.

Now, we prove that N is normal subgroup of G .

Let $a \in G$ and $n \in N$. Then $ana^{-1} = eana^{-1} \in NaNa^{-1} = Naa^{-1} = N$.

$\therefore ana^{-1} \in N$.

Hence N is a normal subgroup of G .

Theorem: 3.15

Let N be a normal subgroup of a group G . Then G/N is a group under the operation defined by $NaNb = Nab$.

Proof:

By theorem 3.14 the operation given by $NaNb = Nab$ is a well defined binary operation in G/N .

Now, let $Na, Nb, Nc \in G/N$.

$$\text{Then } Na(NbNc) = Na(Nbc)$$

$$= Na(bc)$$

$$= N(ab)c$$

$$= (NaNb)Nc.$$

$\therefore$ The binary operation is associative.

$Ne = N \in G/N$ and

$$NaNe = Nae = Na = NeNa.$$

$\therefore Ne$ is the identity element.

$$\text{Also } NaNa^{-1} = Naa^{-1} = Ne = Na^{-1}Na.$$

$\therefore Na^{-1}$ is the inverse of Na .

$\therefore G/N$ is a group.

Definition: 3.2.2

Let N be a normal subgroup of G . Then the group G/N is called **quotient group (factor group)** of G modulo N .

Example:

$3\mathbb{Z}$ is a normal subgroup of $(\mathbb{Z}, +)$. The quotient group $\mathbb{Z}/3\mathbb{Z} = \{3\mathbb{Z} + 0, 3\mathbb{Z} + 1, 3\mathbb{Z} + 2\}$. Hence $\mathbb{Z}/3\mathbb{Z}$ is a group of order 3.

Exercises:

- Find the order of the following quotient groups.

(a) $\mathbb{Z}_6 / \langle 3 \rangle$.

Space for Hints

- (b) $Z_{60} / \langle 5 \rangle$.
2. Compute S_n / A_n .
 3. Compute $V_4 / \{e, a\}$.
 4. Prove that if G is abelian and H is a subgroup of G , then G/H is abelian.
 5. Prove that if G is a cyclic group and H is a subgroup of G , then G/H is cyclic.
 6. Show that $|G/H| = \frac{|G|}{|H|}$ if H is a normal subgroup of a finite group G .
 7. Determine which of the following statements are true and which are false.
 - (a) Z/nZ is cyclic.
 - (b) Quotient group of any cyclic group is cyclic.
 - (c) Quotient group of an abelian group is abelian.
 - (d) S_n / A_n is abelian.
 - (e) The quotient group of a non-abelian group is non-abelian.

Answers:

1. (a) 3 (b) 5
7. (a) T (b) T (c) T (d) T (e) F.

UNIT – 4

HOMOMORPHISMS

4.1 Introduction:

In this section we consider an important concept known as “homomorphism”. The notion of homomorphism is common to all aspects of modern algebra. A homomorphism of group is a map which preserves composition.

Definition: 4.1.1

A map f from a group G into a group G' is called a **homomorphism** if $f(ab) = f(a)f(b)$ for all $a, b \in G$.

Obviously every isomorphism is a homomorphism and a bijective homomorphism is an **isomorphism**.

Examples:

1. $f : (Z, +) \rightarrow (Z, +)$ defined by $f(x) = 2x$ is a homomorphism. For, $f(x + y) = 2(x + y) = 2x + 2y = f(x) + f(y)$. Note that f is 1 – 1.

2. $f : (R^*, \cdot) \rightarrow (R^+, \cdot)$ defined by $f(x) = |x|$ is a homomorphism. For, $f(xy) = |xy| = |x| |y| = f(x)f(y)$. This homomorphism is onto.

3. $f : G \rightarrow G'$ defined by $f(a) = e'$, where e' is the identity in G' is a trivial homomorphism. For, $f(ab) = e' = e'e' = f(a)f(b)$.

4. $f : (Z, +) \rightarrow (C^*, \cdot)$ given by $f(n) = i^n$ is a homomorphism. For, $f(m + n) = i^{n+m} = i^n i^m = f(n)f(m)$. Note that f is neither 1 – 1 nor onto.

5. $f : (R \times R, +) \rightarrow (R, +)$ given by $f(x, y) = x$ is a homomorphism.

6. Let G be a group and N a normal subgroup of G . $f : G \rightarrow G/N$ given by $f(a) = Na$ is a homomorphism.

For, $f(ab) = Nab = NaNb = f(a)f(b)$.

f is called the **canonical homomorphism** from G to G/N . Note that f is onto.

4.2 Types of Homomorphisms:

Definition: 4.2.1

Let $f : G \rightarrow G'$ be a homomorphism.

- (i) If f is onto, then it is called an **epimorphism**.
- (ii) If f is 1 - 1, then it is called a **monomorphism**.

Note:

If $f : G \rightarrow G'$ is an epimorphism then G' is called a **homomorphic image** of G .

A homomorphism of a group to itself is called an **endomorphism**.

Theorem: 4.1

Let $f : G \rightarrow G'$ be a homomorphism. Then

- (i) $f(e) = e'$.
- (ii) $f(a^{-1}) = [f(a)]^{-1}$.
- (iii) If H is a subgroup of G then $f(H)$ is a subgroup of G' .
- (iv) If H is normal in G , then $f(H)$ is normal in $f(G)$.
- (v) If H' is a subgroup of G' , then $f^{-1}(H')$ is a subgroup of G .
- (vi) If H' is normal in $f(G)$ then $f^{-1}(H')$ is normal in G .

Proof:

- (i) Let $a \in G$.

$$\text{Then } f(a) = f(ae) = f(a)f(e).$$

$$\text{Hence } f(e) = e'.$$

- (ii) $f(a)f(a^{-1}) = f(e) = e'.$

$$\text{Hence } f(a^{-1}) = [f(a)]^{-1}.$$

- (iii) Let H be a subgroup of G .

Since H is non - empty, $f(H)$ is also non - empty.

Now, let $x, y \in f(H)$.

Then $x = f(a)$ and $y = f(b)$ where $a, b \in H$.

$$\begin{aligned} \therefore xy^{-1} &= f(a)[f(b)]^{-1} \\ &= f(a)f(b^{-1}) = f(ab^{-1}). \end{aligned}$$

Now, since H is a subgroup of G , $ab^{-1} \in H$.

$$\therefore xy^{-1} = f(ab^{-1}) \in f(H).$$

$$\therefore f(H) \text{ is a subgroup of } G'.$$

- (iv) Let H be normal in G . Let $x \in f(H)$ and $y \in f(G)$.

We claim that $xyx^{-1} \in f(H)$.

Now, $x = f(a)$ and $y = f(b)$ where $a \in H$ and $b \in G$.

Since H is normal in G , $bab^{-1} \in H$.

$$\therefore f(bab^{-1}) \in f(H).$$

$$\therefore f(b)f(a)f(b^{-1}) \in f(H).$$

$$\therefore yxy^{-1} \in f(H). \text{ Hence } f(H) \text{ is normal in } f(G).$$

(v) Since $f(e) = e' \in H'$; $e' \in f^{-1}(H')$ and hence $f^{-1}(H') \neq \Phi$.

Now, let $a, b \in f^{-1}(H')$.

Then $f(a), f(b) \in H'$.

$$\therefore f(a)[f(b)]^{-1} \in H'.$$

$$\therefore f(ab^{-1}) \in H' \text{ (ie), } ab^{-1} \in f^{-1}(H').$$

Hence $f^{-1}(H')$ is a subgroup of G .

(vi) Let $x \in f^{-1}(H')$ and $a \in G$.

Then $f(x) \in H'$ and $f(a) \in f(G)$.

Since H' is normal in $f(G)$, $f(a)f(x)[f(a)]^{-1} \in H'$.

$$\therefore f(axa^{-1}) \in H'.$$

Hence $axa^{-1} \in f^{-1}(H')$.

Thus $f^{-1}(H')$ is normal in G .

Examples:

1. Consider the homomorphism $f : (Z, +) \rightarrow (Z_n, \oplus)$ which is given in the beginning of this section.

Let $K = \{x / x \in Z, f(x) = 0\}$.

Clearly $K = nZ$ which is a normal subgroup of Z .

2. Consider the homomorphism $f : (R^*, \cdot) \rightarrow (R^+, \cdot)$ which is given by $f(x) = |x|$.

Let $K = \{x / x \in R^*, f(x) = 1\}$. Clearly $K = \{1, -1\}$ which is a normal subgroup of $(R^*, \cdot)$.

Definition: 4.2.2

Let $f : G \rightarrow G'$ be a homomorphism. Let

Space for Hints

$K = \{x / x \in G, f(x) = e'\}$. Then K is called the **kernel** of f and is denoted by $\ker f$.

Theorem: 4.2

Let $f: G \rightarrow G'$ be a homomorphism. Then the kernel K of f is a normal subgroup of G .

Proof:

Since $f(e) = e', e \in K$ and hence $K \neq \Phi$.

Now, let $x, y \in K$. Then $f(x) = e' = f(y)$.

$$\therefore f(xy^{-1}) = f(x)f(y^{-1}) = f(x)[f(y)]^{-1} = e'(e')^{-1} = e'e' = e'.$$

Thus $xy^{-1} \in K$. Hence K is a subgroup of G .

Now, let $x \in K$ and $a \in G$.

Then

$$\begin{aligned} f(axa^{-1}) &= f(a)f(x)f(a^{-1}) \\ &= f(a)e'[f(a)]^{-1} \\ &= f(a)[f(a)]^{-1} \\ &= e'. \end{aligned}$$

$\therefore axa^{-1} \in K$. Hence K is a normal subgroup of G .

Aliter:

$\{e'\}$ is a normal subgroup of $f(G)$. Hence $\ker f = f^{-1}(\{e'\})$ is a normal subgroup of G .

Problems:

Problem: 1

Let $f: G \rightarrow G'$ be a homomorphism. Then f is 1-1 iff $\ker f = \{e\}$.

Solution:

Obviously f is 1-1 $\Rightarrow \ker f = \{e\}$.

Conversely, let $\ker f = \{e\}$.

We prove f is 1-1.

$$f(x) = f(y) \Rightarrow f(x)[f(y)]^{-1} = e'.$$

$$\Rightarrow f(xy^{-1}) = e'.$$

$$\Rightarrow xy^{-1} \in \ker f$$

$$\Rightarrow xy^{-1} = e.$$

$$\Rightarrow x = y.$$

Hence f is 1-1.

Problem: 2

Let G be any group and H be the centre of G . Then $G/H \cong I(G)$, the group of inner automorphisms of G .

Solution:

Consider $f: G \rightarrow I(G)$ defined by $f(a) = \phi_a$.

Then $f(ab) = \phi_{ab} = \phi_a \circ \phi_b = f(a)f(b)$.

Hence f is a homomorphism.

Clearly f is onto.

Now, we claim that $\ker f = H$.

$$a \in \ker f \Leftrightarrow f(a) = \phi_e.$$

$$\Leftrightarrow \phi_a = \phi_e$$

$$\Leftrightarrow \phi_a(x) = x \text{ for all } x \in G$$

$$\Leftrightarrow axa^{-1} = x \text{ for all } x \in G$$

$$\Leftrightarrow ax = xa \text{ for all } x \in G$$

$$\Leftrightarrow a \in H.$$

Hence $\ker f = H$.

$\therefore$ By the fundamental theorem of homomorphism $G/H \cong I(G)$.

Problem: 3

Show that $R^*/\{1, -1\} \cong R^+$.

Solution:

Consider $f: R^* \rightarrow R^+$ defined by $f(x) = |x|$.

Clearly f is an epimorphism and $\ker f = \{1, -1\}$.

Hence by the fundamental theorem of homomorphism $R^*/\{1, -1\} \cong R^+$.

Problem: 4

Any homomorphic image of a cyclic group is cyclic.

Solution:

Let G be a cyclic group and $f: G \rightarrow G'$ be an epimorphism. Let a be a generator of G . Then $f(a)$ is a generator of G' . (refer proof of theorem 4.7).

Hence G' is cyclic.

Problem: 5

Show that the map $f: (C, +) \rightarrow (R, +)$ defined by $f(x + iy) = y$ is an epimorphism and $\ker f = R$. Deduce that $C/R \cong R$.

Solution:

Let $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$.

Then $(z_1 + z_2) = (x_1 + x_2) + i(y_1 + y_2)$.

$\therefore f(z_1 + z_2) = y_1 + y_2 = f(z_1) + f(z_2)$.

Hence f is a homomorphism. Clearly f is onto.

Now, $\ker f = \{x + iy / f(x + iy) = 0\}$.

$= \{x + iy / y = 0\}$.

$= R$.

$\therefore$ By the fundamental theorem of homomorphism $C/R \cong R$.

Exercises:

- Determine which of the following maps are homomorphism. If it is a homomorphism, find the kernel.

(a) $f: (Z, +) \rightarrow \{1, -1\}$ given by $f(n) = \begin{cases} 1 & \text{if } n \text{ is even} \\ -1 & \text{if } n \text{ is odd} \end{cases}$.

(b) $f: R^* \rightarrow R^*$ given by $f(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$.

(c) $f: (R \times R, +) \rightarrow (R, +)$ given by $f(x, y) = y$.

(d) $f: (Z, +) \rightarrow (R^*, \cdot)$ given by $f(x) = 3^x$.

(e) $f: S_n \rightarrow \{1, -1\}$ given by

$$f(p) = \begin{cases} 1 & \text{if } p \text{ is an even permutation} \\ -1 & \text{if } p \text{ is an odd permutation} \end{cases}$$

(f) $f: R \rightarrow C$ given by $f(x) = e^{ix}$.

(g) $f: (Z, +) \rightarrow (Z, +)$ given by $f(n) = 2n$.

(h) $f: R^* \rightarrow R^*$ given by $f(x) = -x$.

- (i) $f: \mathbb{Z}_6 \rightarrow \mathbb{Z}_2$ given by $f(x) = \text{remainder of } x \text{ when } x \text{ is divided by } 2$.
- (j) $f: \mathbb{C}^* \rightarrow \mathbb{R}^*$ given by $f(z) = |z|$.
- (k) $f: (\mathbb{R}, +) \rightarrow (\mathbb{R}, +)$ given by $f(x) = x + 2$.
2. Prove that for any positive integer n , $\mathbb{Z}/n\mathbb{Z} \cong \mathbb{Z}_n$.
 3. If $f: G \rightarrow G'$ and $g: G' \rightarrow G''$ are homomorphisms then prove that $g \circ f: G \rightarrow G''$ is a homomorphism.
 4. Show that any homomorphic image of an abelian group is abelian.
 5. Find all the homomorphisms of $(\mathbb{Z}, +)$ onto $(\mathbb{Z}, +)$.
 6. Show that a cyclic group of order 8 is homomorphic to
 - (a) A cyclic group of order 4.
 - (b) A cyclic group of order 2.
 7. Prove that there exist 8 homomorphisms from $(\mathbb{Z}, +)$ to $(\mathbb{Z}_8, \oplus)$. How many of these homomorphisms are onto?
 8. Determine which of the following statements are true and which are false.
 - (a) Any isomorphism is a homomorphism.
 - (b) Any homomorphism is an isomorphism.
 - (c) An infinite group cannot be homomorphic to a finite group.
 - (d) Homomorphism preserves the order of an element.
 - (e) Any homomorphism f is a monomorphism iff $\ker f = \{e\}$.

Answers:

1.

(a) Yes, $\ker f = 2\mathbb{Z}$.(b) Yes, $\ker f = \mathbb{R}^+$.(c) Yes, $\ker f = \mathbb{R} \times \{0\}$.(d) Yes, $\ker f = \{0\}$.(e) Yes, $\ker f = A_n$.

(f) No

(g) Yes, $\ker f = \{0\}$.

(h) No

(i) Yes, $\ker f = \{0, 2, 4\}$.

(j) Yes, $\ker f = \{z \mid z \in \mathbb{C} \text{ and } |z| = 1\}$.

(k) No

8. (a) T (b) F (c) F (d) F (e) T.

4.3 Isomorphism:

4.3 Introduction:

Let $\omega \neq 1$ be a cubic root of unity. Let $G = \{1, \omega, \omega^2\}$. G is a group under usual multiplication. The Cayley table for G is given by

.	1	ω	ω^2
1	1	ω	ω^2
ω	ω	ω^2	1
ω^2	ω^2	1	ω

$(\mathbb{Z}_3, \oplus)$ is a group and its Cayley table is given by

$\oplus$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

We note that these two tables for the groups of order 3 keep the same pattern. In fact any group of order 3 is cyclic and hence it is easily seen that all groups with 3 elements are "like" each other. Thus if two groups G and G' are "like" each other, it should be possible for us to obtain G' from G by renaming each element x in G with the name of an element x' in G' . The remaining of the elements of G can be achieved by means of a bijection $f: G \rightarrow G'$. If $x \in G$ we view $f(x)$ as a new name for x . finally if the groups are to be "like" each other, then if x and y are in G the new name for xy should be $f(x)f(y)$ so that $f(xy) = f(x)f(y)$. Note that the product of xy is computed in G and the product $f(x)f(y)$ is computed in G' . Two groups which are like each other are usually called **isomorphic**. The following definition makes these ideas mathematically precise.

Definition: 4.3.2

Let G and G' be two groups. A map $f: G \rightarrow G'$ is called an **isomorphism** if

- (i) f is a bijection
- (ii) $f(xy) = f(x)f(y)$ for all $x, y \in G$.

Two groups G and G' are said to be **isomorphic** if there exists an **isomorphism** $f: G \rightarrow G'$. If two groups G and G' are isomorphic we write $G \cong G'$.

Theorem: 4.3

Isomorphism is an equivalence relation among groups.

Proof:

For any group G , $i_G: G \rightarrow G$ is clearly an isomorphism.

Hence $G \cong G$. Therefore the relation is reflexive.

Now, let $G \cong G'$ and let $f: G \rightarrow G'$ be an isomorphism.

Then f is a bijection.

$\therefore f^{-1}: G' \rightarrow G$ is also a bijection.

Now, let $x', y' \in G'$.

Let $f^{-1}(x') = x$ and $f^{-1}(y') = y$.

Then $f(x) = x'$ and $f(y) = y'$.

$\therefore f(xy) = f(x)f(y) = x'y'$.

$\therefore f^{-1}(x'y') = xy = f^{-1}(x')f^{-1}(y')$.

Hence f^{-1} is an isomorphism.

Thus $G' \cong G$ and hence the relation is symmetric.

Now, let $G \cong G'$ and $G' \cong G''$.

Then there exist isomorphisms $f: G \rightarrow G'$ and $g: G' \rightarrow G''$.

Since f and g are bijections, $g \circ f: G \rightarrow G''$ is also a bijection.

Now, let $x, y \in G$. Then

$$\begin{aligned} (g \circ f)(xy) &= g[f(xy)] \\ &= g[f(x)f(y)] \text{ (since } f \text{ is an isomorphism)} \\ &= g[f(x)]g[f(y)] \text{ (since } g \text{ is an isomorphism)} \\ &= (g \circ f)(x)(g \circ f)(y). \end{aligned}$$

Hence $g \circ f$ is an isomorphism.

Thus $G \cong G''$ and hence the relation is transitive.

$\therefore$ Isomorphism is an equivalence relation among groups.

Examples:

1. $(\mathbb{Z}, +) \cong (2\mathbb{Z}, +)$.

Consider $f: \mathbb{Z} \rightarrow 2\mathbb{Z}$ given by $f(x) = 2x$.

Clearly f is a bijection. Also

$$\begin{aligned} f(x+y) &= 2(x+y) \\ &= 2x + 2y = f(x) + f(y). \end{aligned}$$

Hence f is an isomorphism.

2. Let $G = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \mid a \in \mathbb{R}^* \right\}$.

G is a group under matrix multiplication.

We claim that $G \cong (\mathbb{R}^*, \cdot)$.

Consider $f: G \rightarrow \mathbb{R}^*$ given by $f\left(\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}\right) = a$.

Clearly f is a bijection.

Now, let $A = \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ and

$$B = \begin{pmatrix} b & 0 \\ 0 & 0 \end{pmatrix} \in G.$$

$$\text{Then } AB = \begin{pmatrix} ab & 0 \\ 0 & 0 \end{pmatrix}$$

$$\therefore f(AB) = ab = f(A)f(B).$$

Hence f is an isomorphism.

3. $(\mathbb{R}, +) \cong (\mathbb{R}^+, \cdot)$.

Consider $f: \mathbb{R} \rightarrow \mathbb{R}^+$ given by $f(x) = e^x$.

Clearly f is a bijection.

$$\text{Also } f(x+y) = e^{x+y} = e^x e^y = f(x)f(y).$$

Hence f is an isomorphism.

4. $G = \mathbb{R} - \{-1\}$ is a group under $*$ defined by $a*b = a + b + ab$. We

claim that $GG \cong (R^*, \cdot)$.

Consider $f: G \rightarrow R^*$ given by $f(x) = x + 1$. Clearly f is a bijection.

Also $f(x * y) = f(x + y + xy)$.

$$= x + y + xy + 1.$$

$$= (x + 1)(y + 1).$$

$$= f(x)f(y).$$

Hence f is an isomorphism.

5. $(Z_n, \oplus)$ is a group.

Let G denote the set of all n^{th} roots of unity. G is a group under usual multiplication. We claim that $(Z_n, \oplus) \cong G$.

Consider $Z_n \rightarrow G$ given by $f(m) = \omega^m$ where $\omega = \cos(2\pi/n) + i \sin(2\pi/n)$.

Clearly f is a bijection.

Let $a, b \in Z_n$. Let $a + b = qn + r$ where $0 \leq r < n$.

Then $a \oplus b = r$. Hence $f(a \oplus b) = \omega^r$ (1)

Also $f(a)f(b) = \omega^a \omega^b = \omega^{a+b} = \omega^{qn+r} = \omega^{qn} \omega^r = 1 \omega^r = \omega^r$ (2)

From (1) and (2), we get $f(a \oplus b) = f(a)f(b)$.

Hence f is an isomorphism.

Exercises:

1. Show that $G = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \mid a \in R \right\}$ is a group under matrix addition and prove that $G \cong (R, +)$.

2. Show that $G = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \mid a, b \in R \text{ and } a^2 + b^2 \neq 0 \right\}$ is a group under matrix multiplication and $G \cong (C^*, \cdot)$.

3. Let $f_a: R \rightarrow R$ be the function defined by $f_a(x) = x + a$. Then $G = \{f_a \mid a \in R\}$ is a group under composition of functions. Show that $G \cong (R, +)$.

Space for Hints

4. $\mathbb{R} \times \mathbb{R}$ is a group under $+$ defined by $(a,b) + (c,d) = (a+c, b+d)$. Show that $(\mathbb{R} \times \mathbb{R}, +) \cong (\mathbb{C}, +)$.
5. Show that $(\mathbb{Z}, +) \cong (n\mathbb{Z}, +)$.
6. $G = \{2^n / n \in \mathbb{Z}\}$ is a group under usual multiplication. Show that $G \cong (\mathbb{Z}, +)$.
7. In $\mathbb{Z}$ we define $a * b = a + b + 1$. $(\mathbb{Z}, *)$ is a group. Show that $(\mathbb{Z}, *) \cong (\mathbb{Z}, +)$.
8. $G = \{a + b\sqrt{2} / a, b \in \mathbb{Z}\}$ is a group under usual addition. $\mathbb{Z} \times \mathbb{Z}$ is a group under addition defined by $(a,b) + (c,d) = (a+c, b+d)$. Show that $G \cong \mathbb{Z} \times \mathbb{Z}$.
9. $\mathbb{R}^*$ is a group under the binary operation $*$ defined by $a * b = \frac{1}{2}ab$. Show that $(\mathbb{R}^*, *) \cong (\mathbb{R}^*, \cdot)$.
10. $G = \{1, i, -1, -i\}$ is a group under multiplication. Show that $f : G \rightarrow (\mathbb{Z}_4, \oplus)$ defined by $f(1) = 0, f(i) = 1, f(-1) = 2, f(-i) = 3$ is an isomorphism.
11. Show that $(\mathcal{P}(\{1,2\}), \Delta)$ is isomorphic to the group of symmetries of a rectangle.

Theorem: 4.4

Let $f : G \rightarrow G'$ be an isomorphism. Then

- (i) $f(e) = e'$ where e and e' are the identity elements of G and G' respectively. (ie). In an isomorphism identity is mapped onto identity.
- (ii) $f(a^{-1}) = [f(a)]^{-1}$.

Proof:

- (i) To prove that $f(e) = e'$ it is enough if we prove that $a'f(e) = f(e)a' = a'$ for all $a' \in G'$.

Let $a' \in G'$. Since $f : G \rightarrow G'$ is a bijection, there exists such that $a \in G$ such that $f(a) = a'$.

$$\therefore a'f(e) = f(a)f(e) = f(ae) = f(a) = a'.$$

Similarly, $f(e)a' = a'$.

$$\therefore f(e) = e'.$$

(ii) It is enough to prove that

$$f(a)f(a^{-1}) = f(a^{-1})f(a) = e'.$$

$$\text{Now, } f(a)f(a^{-1}) = f(aa^{-1}) = f(e) = e'$$

$$\bullet \text{ Also, } f(a^{-1})f(a) = f(a^{-1}a) = f(e) = e'.$$

$$\therefore f(a)f(a^{-1}) = f(a^{-1})f(a) = e'.$$

$$\therefore [f(a)]^{-1} = f(a^{-1}).$$

Remark:

The concept of isomorphism for groups is extremely important. Since two isomorphic groups G and G' have essentially the same structure, if one group G has an additional property (for example abelian or cyclic) then the group G' also has this additional property. This is seen in the following three theorems.

Theorem: 4.5

Let $f: G \rightarrow G'$ be an isomorphism. If G is abelian, then G' is also abelian.

Proof:

Let $a', b' \in G'$. Then there exist $a, b \in G$ such that $f(a) = a'$ and $f(b) = b'$.

$$\text{Now, } a'b' = f(a)f(b) = f(ab) = f(ba) = f(b)f(a) = b'a'.$$

Hence G' is abelian.

Theorem: 4.6

Let $f: G \rightarrow G'$ be an isomorphism. Let $a \in G$. Then the order of a is equal to the order of $f(a)$. (ie) Isomorphism preserves the order of each element in a group.

Proof:

Suppose the order of a is n . Then n is the least positive integer such that $a^n = e$.

Now,

$$[f(a)]^n = f(a) \dots \dots \dots f(a) \quad (f(a) \text{ written } n \text{ times})$$

$$\begin{aligned}
 &= f(a^n) && \text{(since } f \text{ is an isomorphism)} \\
 &= f(e) \\
 &= e'.
 \end{aligned}$$

Now, if possible let m be a positive integer such that $0 < m < n$ and $[f(a)]^m = e'$.

But $f(e) = e'$. Since f is 1-1 we have $a^m = e$ which contradicts the definition of the order of a .

$\therefore$ n is the least positive integer such that $[f(a)]^n = e'$.

$\therefore$ The order of $f(a)$ is n .

Theorem: 4.7

Let $f: G \rightarrow G'$ be an isomorphism. If G is cyclic then G' is also cyclic.

Proof:

Let a be a generator of the group G . We shall prove that $f(a)$ is a generator of the group G' .

Let $x' \in G'$. Since f is a bijection, there exists $x \in G$ such that $f(x) = x'$.

Now, since $G = \langle a \rangle$, $x = a^n$ for some integer n .

Hence $x' = f(x) = f(a^n) = [f(a)]^n$.

Since $x' \in G'$ is arbitrary every element of G' is of the form $[f(a)]^n$ so that $G' = \langle f(a) \rangle$.

Hence G' is cyclic.

Problems:

Problem: 1

Show that $(\mathbb{R}^*, \cdot)$ is not isomorphic to $(\mathbb{R}, +)$.

Solution:

In $(\mathbb{R}, +)$ every element other than 0 is of infinite order. But in $(\mathbb{R}^*, \cdot)$ there exists an element (other than 1) of finite order. For example, -1 is of order 2 in $(\mathbb{R}^*, \cdot)$. Hence we cannot find an isomorphism from $(\mathbb{R}^*, \cdot)$ to $(\mathbb{R}, +)$. (by theorem 4.6).

Problem: 2

Show that $(Z_4, \oplus)$ is not isomorphic to V_4 .

Solution:

In Z_4 , 1 is an element of order 4. But in V_4 every element other than e is of order 2. Hence the two groups are not isomorphic.

This can also be proved by noticing that Z_4 is cyclic and V_4 is not cyclic.

Problem: 3

If G is a group and G' is a set with a binary operation and there exists a one – one mapping f from G onto G' such that $f(ab) = f(a)f(b)$ for all $a, b \in G$ then show that G' is also a group.

Solution:

Let $a, b, c \in G'$.

Since $f : G \rightarrow G'$ is a bijection, there exists $a, b, c \in G$ such that

$$f(a) = a'; f(b) = b'; f(c) = c'.$$

Since G is a group, $(ab)c = a(bc)$.

$$\therefore f[(ab)c] = f[a(bc)].$$

$$\therefore f(ab)f(c) = f(a)f(bc) \text{ (by hypothesis)}$$

$$\therefore [f(a)f(b)]f(c) = f(a)[f(b)f(c)].$$

$$\therefore (a'b')c' = a'(b'c').$$

$\therefore$ The binary operation in G' is associative.

Now, let $e \in G$ be the identity element.

Let $a' \in G'$. Since $f : G \rightarrow G'$ is a bijection, there exists $a \in G$ such that $f(a) = a'$.

Now, $ae = ea = a$.

$$\therefore f(ae) = f(ea) = f(a).$$

$$\therefore f(a)f(e) = f(e)f(a) = f(a).$$

$$\therefore a'f(e) = f(e)a' = a'.$$

$$\therefore f(e) \text{ is the identity in } G'.$$

Let $a' \in G'$. Since $f : G \rightarrow G'$ is a bijection, there exists $a \in G$ such that $f(a) = a'$.

Space for Hints

Now, $aa^{-1} = a^{-1}a = e$.

$$\therefore f(aa^{-1}) = f(a^{-1}a) = f(e).$$

$$\therefore f(a)f(a^{-1}) = f(a^{-1})f(a) = f(e).$$

$$\therefore a'f(a^{-1}) = f(a^{-1})a' = f(e).$$

$$\therefore f(a^{-1}) \text{ is the inverse of } a' \text{ in } G'.$$

Hence G' is a group.

Problem: 4

Let G be any group. Show that $f: G \rightarrow G$ given by $f(x) = x^{-1}$ is an isomorphism $\Leftrightarrow G$ is abelian.

Solution:

Let $f: G \rightarrow G$ given by $f(x) = x^{-1}$ be an isomorphism.

We claim that G is abelian.

Let $x, y \in G$.

Then $f(x^{-1}y^{-1}) = f(x^{-1})f(y^{-1})$. (since f is an isomorphism).

$$\therefore (x^{-1}y^{-1})^{-1} = (x^{-1})^{-1}(y^{-1})^{-1}.$$

$$\therefore (y^{-1})^{-1}(x^{-1})^{-1} = (x^{-1})^{-1}(y^{-1})^{-1}.$$

$$\therefore yx = xy.$$

Hence G is abelian.

Conversely, suppose G is abelian.

Clearly $f: G \rightarrow G$ given by $f(x) = x^{-1}$ is a bijection.

Now,

$$\begin{aligned} f(xy) &= (xy)^{-1} \\ &= y^{-1}x^{-1} \\ &= x^{-1}y^{-1} \text{ (since } G \text{ is abelian)} \\ &= f(x)f(y). \end{aligned}$$

$\therefore f$ is isomorphism.

Exercises:

1. Show that any two groups of order 2 are isomorphic.
2. Show that any two groups of order 3 are isomorphic.
3. Show that any proper subgroup of $(\mathbb{Z}, +)$ is isomorphic to $(\mathbb{Z}, +)$.

(Hint: Any proper subgroup of $(\mathbb{Z}, +)$ is $n\mathbb{Z}$).

4. Show that $(\mathbb{Q}, +)$ is not isomorphic to $(\mathbb{Q}^*, \cdot)$.
5. Show that $(\mathbb{C}, +)$ is not isomorphic to $(\mathbb{C}^*, \cdot)$.
6. Let $f : G \rightarrow G'$ be an isomorphism. Then if H is a subgroup of G , $f(H)$ is a subgroup of G' .
7. Prove that for any positive integer n , $(\mathbb{Z}/n\mathbb{Z}) \cong \mathbb{Z}_n$.

Theorem: 4.8

Any infinite cyclic group G is isomorphic to $(\mathbb{Z}, +)$

Proof:

Let G be an infinite cyclic group with generator a . Then $G = \{a^n / n \in \mathbb{Z}\}$.

Define $f : \mathbb{Z} \rightarrow G$ by $f(n) = a^n$.

Since G is infinite, $n \neq m \Rightarrow a^n \neq a^m$.

Hence f is 1-1. Obviously f is onto.

Now, $f(n+m) = a^{n+m} = a^n a^m = f(n)f(m)$.

Hence f is an isomorphism.

Corollary:

Any two infinite cyclic groups are isomorphic to each other.

Let G and G' be two infinite cyclic groups. By theorem 4.8, $G \cong G'$
(since $\cong$ is an equivalence relation).

Theorem: 4.9

Any finite cyclic group of order n is isomorphic to $(\mathbb{Z}_n, \oplus)$.

Proof:

Let G be a cyclic group of order n with generator a . Then $G = \{e, a, a^2, \dots, a^{n-1}\}$.

Define $f : \mathbb{Z}_n \rightarrow G$ by $f(r) = a^r$.

Clearly f is a bijection.

Now, let $r, s \in \mathbb{Z}_n$. Let $r \oplus s = t$.

Then $r + s = qn + t$, where $0 \leq t < n$.

$$\therefore f(r \oplus s) = a^{r \oplus s} = a^t \quad (1)$$

Space for Hints

Also,

$$\begin{aligned} f(r)f(s) &= a^r a^s = a^{r+s} = a^{qn+t} = a^{qn} a^t \\ &= (a^n)^q a^t = e a^t = a^t \end{aligned} \quad (2)$$

From (1) and (2), we get $f(r \oplus s) = f(r)f(s)$.

Hence f is an isomorphism.

Corollary:

Any two finite cyclic groups of the same order are isomorphic.

CAYLEY'S THEOREM

Theorem: 5.1 (Cayley's theorem)**Statement:**

Any finite group is isomorphic to a group of permutations.

Proof:

We shall prove this theorem in 3 steps. We shall first find a set G' of permutations. Then we prove that G' is a group of permutations and finally we exhibit an isomorphism $\phi: G \rightarrow G'$.

Step: 1

Let G be a finite group of order n . Let $a \in G$.

Define $f_a: G \rightarrow G$ by $f_a(x) = ax$.

Now, f_a is 1-1, since $f_a(x) = f_a(y) \Rightarrow ax = ay \Rightarrow x = y$.

f_a is onto (since, if $y \in G$, then $f_a(a^{-1}y) = a(a^{-1}y) = y$).

Thus f_a is a bijection.

Since G has n elements, f_a is just a permutation on n symbols.

Let $G' = \{f_a / a \in G\}$.

Step: 2

We prove G' is a group. Let $f_a, f_b \in G'$.

$$(f_a \circ f_b)(x) = f_a(f_b(x)) = f_a(bx) = a(bx) = (ab)x = f_{ab}(x).$$

Hence $f_a \circ f_b = f_{ab}$. Hence G' is closed under composition of mappings. $f_e \in G'$ is the identity element.

The inverse of f_a in G' is f_a^{-1} .

Step: 3

We prove $G \cong G'$.

Define $\phi: G \rightarrow G'$ by $\phi(a) = f_a$.

$$\phi(a) = \phi(b) \Rightarrow f_a = f_b \Rightarrow f_a(x) = f_b(x) \Rightarrow ax = bx \Rightarrow a = b.$$

Hence ϕ is 1-1. Obviously ϕ is onto.

$$\text{Also } \phi(ab) = f_{ab} = f_a \circ f_b = \phi(a) \circ \phi(b).$$

Hence ϕ is an isomorphism.

Example:

Consider the group $G = \{e, a, b\}$ whose multiplication table is given by

$\cdot$	e	a	b
e	e	a	b
a	a	b	e
b	b	e	a

By Cayle's theorem G is isomorphic to the permutation group $G' = \{f_e, f_a, f_b\}$ where

$$f_e = \begin{pmatrix} e & a & b \\ e & a & b \end{pmatrix}; f_a = \begin{pmatrix} e & a & b \\ a & b & e \end{pmatrix} \text{ and } f_b = \begin{pmatrix} e & a & b \\ b & e & a \end{pmatrix}$$

Definition: 5.1.1

An isomorphism of a group G to itself called an **automorphism** of G .

Examples:

1. Any group G has atleast one automorphism namely i_G .
2. The map $f: R^* \rightarrow R^*$ defined by $f(a) = a^{-1}$ is an automorphism. Clearly f is a bijection.

$$\text{Also } f(ab) = (ab)^{-1} = b^{-1}a^{-1} = a^{-1}b^{-1} = f(a)f(b).$$

More generally if G is abelian, $f: G \rightarrow G$ defined by $f(a) = a^{-1}$ is an automorphism.

3. The mapping ϕ given by $\phi(z) = \bar{z}$ is an automorphism of the additive group of complex numbers. Clearly ϕ is a bijection and

$$\phi(z + \omega) = \overline{(z + \omega)}.$$

$$= \bar{z} + \bar{\omega}$$

$$= \phi(z) + \phi(\omega).$$

4. Let G be any group. Let $a \in G$. Then $\phi_a: G \rightarrow G$ defined by $\phi_a(x) = axa^{-1}$ is a automorphism of G .

For, let $x, y \in G$. Then $\phi_a(x) = \phi_a(y) \Rightarrow axa^{-1} = aya^{-1} \Rightarrow x = y$ (by

cancellation law)

$$\therefore \phi_a \text{ is } 1 - 1.$$

$$\text{Also } \phi_a(a^{-1}xa) = a(a^{-1}xa)a^{-1}$$

$$= (aa^{-1})x(aa^{-1}) = exe = x.$$

Hence $a^{-1}xa$ is the pre - image of x under ϕ_a . Also

$$\phi_a(xy) = axya^{-1}$$

$$= (axa^{-1})(aya^{-1})$$

$$= \phi_a(x)\phi_a(y).$$

Thus ϕ_a is an automorphism of G .

Definition: 5.1.2

The automorphism $\phi_a : G \rightarrow G$ defined in example 4 is called an **inner automorphism** of the group G .

Let G be a group. The set of all automorphisms of G is denoted by $\text{Aut } G$. The set of all inner automorphisms of G is denoted by $I(G)$.

Theorem: 5.2

For any group G ,

- (i) $\text{Aut } G$ is a group under composition of functions.
- (ii) $I(G)$ is a normal subgroup of $\text{Aut } G$.

Proof:

- (i) Let $f, g \in \text{Aut } G$.

$\therefore f$ and g are isomorphisms of G to itself.

$\therefore f \circ g$ is an isomorphism of G to itself (Theorem 4.3).

$\therefore f \circ g \in \text{Aut } G$.

$f \in \text{Aut } G \Rightarrow f^{-1} \in \text{Aut } G$ (Theorem 4.3)

Clearly composition of functions is associative.

Hence $\text{Aut } G$ is a group.

- (ii) Let $\phi_a, \phi_b \in I(G)$. Then

$$(\phi_a \phi_b)(x) = \phi_a(bxb^{-1})$$

$$= a(bxb^{-1})a^{-1}$$

$$= (ab)x(ab)^{-1}$$

$$= \phi_{ab}(x).$$

Hence $\phi_a \phi_b = \phi_{ab} \in I(G)$.

ϕ_e is the identity element of $I(G)$ and inverse of ϕ_a is $\phi_{a^{-1}}$.

$\therefore I(G)$ is a subgroup of $\text{Aut } G$.

We now prove that $I(G)$ is a normal subgroup of $\text{Aut } G$.

Let $\alpha \in \text{Aut}(G)$ and $\phi_a \in I(G)$. Then

$$(\alpha \phi_a \alpha^{-1})(x) = \alpha \phi_a (\alpha^{-1}(x))$$

$$= \alpha(a \alpha^{-1}(x) a^{-1})$$

$$= \alpha(a) \alpha \alpha^{-1}(x) \alpha(a^{-1})$$

$$= \alpha(a) x [\alpha(a)]^{-1}$$

$$= \phi_{\alpha(a)}(x).$$

$$\therefore \alpha \phi_a \alpha^{-1} = \phi_{\alpha(a)} \in I(G).$$

Hence $I(G)$ is a normal subgroup of $\text{Aut } G$.

Theorem: 5.3

Let G be a cyclic group generated by a . Let $f: G \rightarrow G$ be a mapping such that $f(xy) = f(x)f(y)$. Then f is an automorphism of G iff $f(a)$ is a generator of G .

Proof:

Let f be an automorphism of G . We shall prove that $f(a)$ is a generator of G .

Case: (i)

Let G be a finite cyclic group of order n . Then order of a is n . By theorem 4.5 $f(a)$ is also an element of order n and hence $f(a)$ is a generator of G .

Case: (ii)

Let G be infinite. Suppose $f(a)$ is not a generator of G . Let $H = \langle f(a) \rangle$. Then H is a proper subgroup of G .

We claim that $f(G) = H$.

Let $x' \in f(G)$. Then $x' = f(x)$ for some $x \in G$.

Now, $x = a^n$ for some n since $G = \langle a \rangle$.

$$\therefore x = f(a^n) = [f(a)]^n \in H.$$

$$\therefore f(G) \subseteq H.$$

Now, let $x \in H$. Then $x = [f(a)]^n$ for some n .

$$\therefore x = f(a^n). \text{ Hence } x \in f(G).$$

$$\therefore H \subseteq f(G). \text{ Hence } f(G) = H.$$

Since H is a proper subgroup of G , f is not onto which is a contradiction. Hence $f(a)$ is a generator of G .

Conversely let $f: G \rightarrow G$ be a mapping such that $f(xy) = f(x)f(y)$ and let $f(a)$ be a generator of G . We shall prove that f is an automorphism.

It is enough if we prove that f is 1-1 and onto.

Let $x \in G$. Since $f(a)$ is a generator of G , $x = [f(a)]^n$ for some n .

Clearly $f(a^n) = [f(a)]^n = x$. Thus x has a pre image a^n under f .

Hence f is onto.

Now, to prove f is 1-1.

Case (i) G is finite.

Since any function from a finite set onto itself is necessarily 1-1 (verify), f is 1-1.

Case (ii) G is infinite.

Let $x, y \in G$ and let $x = a^n, y = a^m$ and $n \geq m$.

Now,

$$f(x) = f(y) \Rightarrow f(a^n) = f(a^m)$$

$$\Rightarrow [f(a)]^n = [f(a)]^m$$

$$\Rightarrow [f(a)]^{n-m} = e$$

$$\Rightarrow n - m = 0$$

(since $f(a)$ is an element of finite order)

$$\Rightarrow n = m$$

$$\Rightarrow a^n = a^m$$

$$\Rightarrow x = y.$$

Space for Hints

Hence f is 1-1. Thus f is an automorphism.

Note:

Let G be a cyclic group generated by a . Then any automorphism $f: G \rightarrow G$ is completely determined by the image $f(a)$ of the generator. For, if $x \in G$ is any element then $x = a^n$ for some integer n and hence $f(x) = f(a^n) = [f(a)]^n$.

As an example, consider $(Z_4, \oplus)$. Here 1 is a generator of this cyclic group.

If $f(1) = 3$, then

$$f(2) = f(1 \oplus 1) = f(1) \oplus f(1) = 3 \oplus 3 = 2;$$

$$f(3) = f(2 \oplus 1) = f(2) \oplus f(1) = 2 \oplus 3 = 1 \text{ and}$$

$$f(0) = f(3 \oplus 1) = f(3) \oplus f(1) = 1 \oplus 3 = 0.$$

Theorem: 5.4

The number of automorphisms of a cyclic group of order n is $\phi(n)$.

Proof:

Let G be a cyclic group of order n . Let $a \in G$ be a generator. If $f: G \rightarrow G$ is an automorphism then f is completely determined by specifying the image of a . The only possible images of a are any one of the generators of G . Hence the number of automorphism is equal to the number of generators of G . But the number of generators of a cyclic group of order n is $\phi(n)$. (by corollary 3 of Theorem 2.24). Hence the number of automorphisms of a cyclic group of order n is $\phi(n)$.

Problems:

Problem: 1

Construct the group of automorphisms of $(Z_4, \oplus)$.

Solution:

1 and 3 are the only 2 generators of Z_4 . Hence there are only 2 automorphisms of Z_4 , say f and g . They are given by $f(1) = 1$ and $g(1) = 3$.

$$\text{Hence } \text{Aut } G = \{f, g\} \cong Z_2.$$

Problem: 2

Construct the group of automorphisms of $(Z, +)$.

Solution:

1 and -1 are the only 2 generators of $\mathbb{Z}$. Hence there are only 2 automorphisms of $\mathbb{Z}$ say f and g . They are given by $f(1) = 1$ and $g(1) = -1$.

1. $f(1) = 1$ gives the identity automorphism. $g(1) = -1$ determines the automorphism given by $g(x) = -x$.

Hence $\text{Aut } \mathbb{Z} = \{f, g\} \cong \mathbb{Z}_2$.

Problem: 3

Let G be a finite abelian group of order n and let m be a positive integer relatively prime n . Then $f: G \rightarrow G$ defined by $f(x) = x^m$ is an automorphism of G .

Solution:

Since m and n are relatively prime, there exist integers u and v such that $mu + nv = 1$.

Now, let $x \in G$.

Then $x = x^{mu+nv} = x^{mu} x^{nv} = x^{mu} e = x^{mu}$.

Hence $x = x^{mu}$.

Now,

$$f(x) = f(y) \Rightarrow x^m = y^m.$$

$$\Rightarrow x^{mu} = y^{mu}$$

$$\Rightarrow x = y.$$

Hence f is 1-1.

Also $f(x^u) = x^{mu} = x$.

$\therefore$ Every element x has pre-image x^u under f . Hence f is onto.

Also $f(xy) = (xy)^m$

$$= x^m y^m \text{ (since } G \text{ is abelian)}$$

$$= f(x)f(y)$$

Hence f is an isomorphism.

Problem: 4

Show that $\text{Aut } \mathbb{Z}_8 \cong V_4$.

Solution:

The generators of Z_8 are 1, 3, 5, 7. The four different automorphisms of Z_8 are f_1, f_2, f_3, f_4 given by

$$f_1(1) = 1; f_2(1) = 3; f_3(1) = 5; f_4(1) = 7.$$

We shall now compute $f_2 \circ f_3$.

$$\begin{aligned} (f_2 \circ f_3)(1) &= f_2(f_3(1)) = f_2(5) \\ &= f_2(1 \oplus 1 \oplus 1 \oplus 1 \oplus 1) \\ &= f_2(1) \oplus f_2(1) \oplus f_2(1) \oplus f_2(1) \oplus f_2(1) \\ &= 3 \oplus 3 \oplus 3 \oplus 3 \oplus 3 \\ &= 7 = f_4(1). \end{aligned}$$

Thus $f_2 \circ f_3 = f_4$.

Similarly we can find $f_i \circ f_j; i, j = 1, 2, 3, 4$. The Cayley table of

$\text{Aut } Z_8$ is

$\circ$	f_1	f_2	f_3	f_4
f_1	f_1	f_2	f_3	f_4
f_2	f_2	f_1	f_4	f_3
f_3	f_3	f_4	f_1	f_2
f_4	f_4	f_3	f_2	f_1

Clearly $\text{Aut } Z_8 \cong V_4$.

Exercises:

1. Compute the group of automorphisms of $(Z_{12}, \oplus)$ and show that $\text{Aut } Z_{12} \cong \text{Aut } Z_8$.
2. Show that in a group G every inner automorphism is identity iff G is abelian.
3. Prove that a subgroup H of G is a normal subgroup iff $\phi_a(H) = H$ for every inner automorphism ϕ_a .
4. Represent Z_4 and V_4 as groups of permutations.

5.2 Fundamental theorem of Homomorphism

Theorem: 5.5 (Fundamental theorem of Homomorphism)

Statement:

Let $f : G \rightarrow G'$ be an epimorphism. Let K be the kernel of f . Then $G/K \cong G'$.

Proof:

Define $\phi : G/K \rightarrow G'$ by $\phi(ka) = f(a)$.

Step (i) ϕ is well defined.

Let $Kb = Ka$. Then $b \in Ka$.

Hence $b = ka$ where $k \in K$.

Now, $f(b) = f(ka) = f(k)f(a) = e'f(a) = f(a)$.

$\therefore \phi(Kb) = f(b) = f(a) = \phi(Ka)$.

Hence $\phi(Ka) = \phi(Kb)$.

Step (ii) ϕ is 1-1.

For $\phi(Ka) = \phi(Kb) \Rightarrow f(a) = f(b)$.

$$\Rightarrow f(a)[f(b)]^{-1} = e'.$$

$$\Rightarrow f(ab^{-1}) = e'.$$

$$\Rightarrow ab^{-1} \in K.$$

$$\Rightarrow a \in Kb.$$

$$\Rightarrow Ka = Kb.$$

Step (iii) ϕ is onto.

Let $a' \in G'$.

Since f is onto, there exists $a \in G$ such that $f(a) = a'$.

Hence $\phi(Ka) = f(a) = a'$.

Step (iv) ϕ is a homomorphism.

$$\phi(KaKb) = \phi(Kab) = f(ab) = f(a)f(b)$$

$$= \phi(Ka)\phi(Kb).$$

Thus ϕ is an isomorphism from G/K onto G' .

$\therefore G/K \cong G'$.

5.3 Rings

5.3.0 Definition and Examples:

A group is an algebraic system with one binary operation. The familiar examples of real numbers and 2×2 matrices are systems which involve two binary operations. In this chapter we study algebraic systems with two binary operations. We start considering the system Z of integers. Z has two binary operations “+” and “.” $(Z, +)$ is an abelian group. Multiplication is an associative binary operation in Z . These two binary operations are connected by the two distributive laws given by $a(b + c) = ab + ac$ and $(a + b)c = ac + bc$. A generalisation of these basic properties in Z leads us to the concept of a system called **ring**.

Definition: 5.3.1

A nonempty set R together with two binary operations denoted by “+” and “.” and called addition and multiplication which satisfy the following axioms is called a **ring**.

- (i) $(R, +)$ is an abelian group.
- (ii) “.” is an associative binary operation on R .
- (iii) $a.(b + c) = a.b + a.c$ and $(a + b).c = a.c + b.c$ for all $a, b, c \in R$.

Notation:

The unique identity of the additive group $(R, +)$ is denoted by 0 and is called the **zero element** of the ring and the unique additive inverse of a is denoted by $-a$.

Examples:

1. $(Z, +, .); (Q, +, .); (R, +, .); (C, +, .)$ are all rings.
2. $(2Z, +, .)$ is a ring.
3. Let $R = \{a + b\sqrt{2} / a, b \in Z\}$.

Clearly R is an abelian group under usual addition.

Let $a + b\sqrt{2}$ and $c + d\sqrt{2} \in R$. Then

$$(a + b\sqrt{2})(c + d\sqrt{2}) = (ac + 2bd) + (bc + ad)\sqrt{2} \in R.$$

Since the two binary operations are the usual addition and multiplication, the distributive laws and the associative law hold.

Thus R is a ring with usual addition and multiplication.

4. Let $R = \{a + ib / a, b \in \mathbb{Z}\}$. Then R is a ring under usual addition and multiplication. This ring is called the **ring of Gaussian integers**. In general, any subset of complex numbers which is a group under addition and is closed for multiplication is a ring (Verify).
5. $\{0\}$ with binary operation '+' and '.' Defined as $0 + 0 = 0$ and $0.0 = 0$ is a ring. This is called the **null ring**.
6. In $R \times R$ we define $(a, b) + (c, d) = (a + c, b + d)$ and $(a, b) \cdot (c, d) = (ac, bd)$.

Here $(R \times R, +)$ is an abelian group. The identity is $(0, 0)$ and the inverse of (a, b) is $(-a, -b)$.

$$\begin{aligned} \text{Further } (a, b) [c, d) + (e, f)] &= (a, b)(c + e, d + f) \\ &= (ac + ae, bd + bf) \\ &= (ac, bd) + (ae, bf) \\ &= (a, b)(c, d) + (a, b)(e, f). \end{aligned}$$

Similarly $[(a, b) + (c, d)](e, f) = (a, b)(e, f) + (c, d)(e, f)$.

Hence $(R \times R, +, \cdot)$ is a ring.

7. Let $(R, +)$ be any abelian group with identity 0.

We define multiplication in R by $ab = 0$ for all $a, b \in R$.

Clearly $a(bc) = 0 = (ab)c$ so that multiplication is associative.

Also $a(b + c) = 0 = ab + ac$ and $(a + b)c = 0 = ac + bc$.

Hence R is a ring under these operations. This ring is called the **zero ring**.

This example shows that any abelian group with identity 0 can be made into a ring by defining $ab = 0$.

8. $(\mathbb{Z}_n, \oplus, \odot)$ is a ring, for, we know that $(\mathbb{Z}_n, \oplus)$ is an abelian group and $\odot$ is an associative binary operation.

We now prove the distributive laws.

Let $a, b, c \in \mathbb{Z}_n$.

Then $b \oplus c \equiv (b + c)(\text{mod } n)$.

Hence $a \odot (b \oplus c) \equiv a(b + c)(\text{mod } n)$.

Also $a \odot b \equiv ab(\text{mod } n)$ and $a \odot c \equiv ac(\text{mod } n)$ so that

$$(a \odot b) \oplus (a \odot c) \equiv (ab + ac)(\text{mod } n).$$

Since $a \odot (b \oplus c)$ and $(a \odot b) \oplus (a \odot c) \in \mathbb{Z}_n$, we have

$$a \odot (b \oplus c) = (a \odot b) \oplus (a \odot c). \text{ Similarly}$$

$$(a \oplus b) \odot c = (a \odot c) \oplus (b \odot c).$$

Hence $(\mathbb{Z}_n, \oplus, \odot)$ is a ring.

9. $(\wp(S), \Delta, \cap)$ is a ring. We know that $(\wp(S), \Delta)$ is an abelian group

Here $\cap$ is an associative binary operation on $\wp(S)$.

It can easily be verified that $A \cap (B \Delta C) = (A \cap B) \Delta (A \cap C)$ and

$$(A \Delta B) \cap C = (A \cap C) \Delta (B \cap C). \text{ Hence } (\wp(S), \Delta, \cap) \text{ is a ring.}$$

10. $M_2(\mathbb{R})$ under matrix addition and multiplication is a ring.

11. Let R be the set of all real functions. We define addition and multiplication by

$$(f + g)(x) = f(x) + g(x) \text{ and } (fg)(x) = f(x)g(x). \text{ Then } R \text{ is a ring.}$$

Clearly addition of functions is associative and commutative.

The constant function 0 defined by $0(x) = 0$ is the zero element of R and $-f$ is the additive inverse of f .

Hence R is an abelian group.

The associativity of multiplication and the distributive laws are consequences of the corresponding properties in $\mathbb{R}$. Hence R is a ring.

12. Let A be any abelian group. Let $\text{Hom}(A)$ be the set of all endomorphisms of A .

Let $f, g \in \text{Hom}(A)$. We define $f + g$ by $(f + g)(x) = f(x) + g(x)$ and $fg = f \circ g$. Then $\text{Hom}(A)$ is a ring.

Proof:

Let $f, g \in \text{Hom}(A)$.

Then $(f + g)(x + y)$

$$= f(x + y) + g(x + y)$$

$$= f(x) + f(y) + g(x) + g(y)$$

$$= f(x) + g(x) + f(y) + g(y)$$

$$= (f + g)(x) + (f + g)(y).$$

Hence $f + g \in \text{Hom}(A)$.

Obviously $+$ is associative.

Since A is an abelian group $f + g = g + f$.

If 0 is the identity element of the group A then the homomorphism 0 defined by $0(a) = 0$, for all $a \in A$ is the zero element of $\text{Hom}(A)$. Now, let $f \in \text{Hom}(A)$. The function $-f$ defined by $(-f)(x) = -[f(x)]$ is also a homomorphism, since

$$(-f)(x + y) = -[f(x + y)]$$

$$= -[f(x) + f(y)]$$

$$= (-f)(x) + (-f)(y).$$

Clearly $f + (-f) = 0$ and hence $-f$ is the additive inverse of f .

Thus $\text{Hom}(A)$ is an abelian group.

Now, $(f \circ g)(x + y)$

$$= f[g(x + y)]$$

$$= f[g(x) + g(y)]$$

$$= f[g(x)] + f[g(y)]$$

$$= (f \circ g)(x) + (f \circ g)(y).$$

Hence $f \circ g \in \text{Hom}(A)$.

Similarly $(f + g) \circ h = f \circ h + g \circ h$.

Thus $\text{Hom}(A)$ is a ring.

13. Let Q be the set of all symbols of the form $a_0 + a_1i + a_2j + a_3k$ where $a_0, a_1, a_2, a_3 \in R$. Two such symbols $a_0 + a_1i + a_2j + a_3k$ and $b_0 + b_1i + b_2j + b_3k$ are defined to be equal iff $a_i = b_i, i = 0, 1, 2, 3$. We now make Q into a ring by defining $+$ and $\cdot$ as follows.

For any

$$x = a_0 + a_1i + a_2j + a_3k \quad \text{and}$$

$$y = b_0 + b_1i + b_2j + b_3k.$$

Space for Hints

$$x + y = (a_0 + a_1i + a_2j + a_3k) + (b_0 + b_1i + b_2j + b_3k)$$

$$= (a_0 + b_0) + (a_1 + b_1)i + (a_2 + b_2)j + (a_3 + b_3)k \text{ and}$$

$$xy = (a_0 + a_1i + a_2j + a_3k)(b_0 + b_1i + b_2j + b_3k)$$

$$= (a_0b_0 - a_1b_1 - a_2b_2 - a_3b_3) + (a_0b_1 + a_1b_0 + a_2b_3 - a_3b_2)i$$

$$+ (a_0b_2 + a_2b_0 + a_3b_1 - a_1b_3)j + (a_0b_3 + a_3b_0 + a_1b_2 - a_2b_1)k.$$

The formula for the product comes from multiplying the two symbols formally and collecting the terms using the relations.

$$i^2 = j^2 = k^2 = ijk = -1, ij = -ji = k, jk = -kj = i \text{ and } ki = -ik = j.$$

Clearly $+$ is associative and commutative. $0 = 0 + 0i + 0j + 0k$ is the zero element.

$-a_0 - a_1i - a_2j - a_3k$ is the additive inverse of $a_0 + a_1i + a_2j + a_3k$.

The associative law of multiplication and the two distributive laws can be easily verified. Hence $(Q, +, \cdot)$ is a ring.

This ring is called the **ring of quaternions**.

14. The set R of all matrices of the form $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$ where $a, b \in R$ is a ring under matrix addition and matrix multiplication.

Proof:

$$\text{Let } A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \text{ and } B = \begin{pmatrix} c & d \\ -d & c \end{pmatrix} \in R.$$

Then

$$A + B = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} + \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} a+c & b+d \\ -(b+d) & a+c \end{pmatrix} \in R.$$

$$AB = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} c & d \\ -d & c \end{pmatrix} = \begin{pmatrix} ac-bd & ad+bc \\ -(ad+bc) & ac-bd \end{pmatrix} \in R.$$

Clearly matrix addition is commutative and associative.

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \in R \text{ is the zero element.}$$

$$\begin{pmatrix} -a & -b \\ b & -a \end{pmatrix} \text{ is the inverse of the matrix } \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

Further matrix multiplication is associative and the distributive laws are valid for

2 x 2 matrices.

Hence R is a ring.

Exercises:

1. Prove that the set of all real numbers of the form $a + b\sqrt{3}$ where $a, b \in \mathbb{Q}$ under usual addition and multiplication is a ring.

2. Prove that the set of all matrices of the form $\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix}$ where $a \in R$ is a ring under matrix addition and multiplication.

3. Show that $(\mathbb{Z}, \oplus, \odot)$ is a ring where $a \oplus b = a + b - 1$ and $a \odot b = a + b - ab$.

4. Show that $(n\mathbb{Z}, +, \cdot)$ is a ring.

5. Verify whether the sets $\{0, 1\}$ and $\{a, b\}$ are rings with operations defined by the following tables.

(a)

+	0	1
0	0	1
1	1	0

.	0	1
0	0	0
1	0	1

(b)

+	a	b
a	a	b
b	b	a

.	a	b
a	a	b
b	a	b

6. Construct the Cayley tables for the ring $(\wp(S), \Delta, \cap)$ when $S = \{1, 2\}$.

7. Show that $(2\mathbb{Z}, +, *)$ is a ring where $+$ is usual addition and $*$ is given by $a * b = \frac{1}{2}ab$.

8. In $\mathbb{R} \times \mathbb{R}$ define $(a, b) + (c, d) = (a + c, b + d)$ and $(a, b)(c, d) = (ac - bd, bc + ad)$. Show that $(\mathbb{R} \times \mathbb{R}, +, \cdot)$ is a ring.

9. Let $S = \{iy / y \in \mathbb{R}\}$. Is S a ring with usual addition and multiplication?

10. Is $(\wp(S), \cup, \cap)$ a ring?

11. Is $(\wp(S), \Delta, \cup)$ a ring?

12. In $\mathbb{Q}$ we define $a \oplus b = ab$ and $a \odot b = a + b$. Show that $(\mathbb{Q}, \oplus, \odot)$ is not a ring.

13. Determine which of the following statements are true and which are false.

- (a) The set of all even integers is a ring under usual addition and multiplication.
- (b) The set of all odd integers is a ring under usual addition and multiplication.
- (c) In any ring addition is commutative.
- (d) The non-zero elements of a ring form a group under multiplication.

Answers:

5. (a) Ring (b) Not a ring 9. No 10. No
11. No
13. (a) T (b) T (c) T (d) F

5.4 Elementary properties of rings:

Theorem: 5.6

Let R be a ring and $a, b \in R$. Then

- (i) $0a = a0 = 0$
- (ii) $a(-b) = (-a)b = -(ab)$
- (iii) $(-a)(-b) = ab$
- (iv) $a(b - c) = ab - ac$.

Proof:

$$(i) \quad a0 = a(0 + 0) = a0 + a0.$$

$\therefore a0 = 0$. (by cancellation law in $(R, +)$) Similarly $0a = 0$.

$$(ii) \quad a(-b) + ab = a(-b + b) = a0 = 0.$$

$$\therefore a(-b) = -(ab).$$

Similarly, $(-a)b = -(ab)$.

$$(iii) \text{ By (ii), } (-a)(-b) = -[a(-b)] = -(-(ab)) = ab.$$

$$(iv) \quad a(b - c) = a[b + (-c)] = ab + a(-c) = ab - ac.$$

Problems:

Problem: 1

If R is a ring such that $a^2 = a$ for all $a \in R$, prove that

- (i) $a + a = 0$

$$(ii) \quad a + b = 0 \Rightarrow a = b$$

$$(iii) \quad ab = ba.$$

Proof:

$$\begin{aligned} (i) \quad a + a &= (a + a)(a + a) \\ &= a(a + a) + a(a + a) \\ &= aa + aa + aa + aa \\ &= (a + a) + (a + a) \text{ (since } a^2 = a) \end{aligned}$$

$$\text{Hence } a + a = 0.$$

$$(ii) \quad \text{Let } a + b = 0. \text{ By (i) } a + a = 0.$$

$$\therefore a + b = a + a \text{ so that } a = b.$$

$$\begin{aligned} (iii) \quad a + b &= (a + b)(a + b) \\ &= a(a + b) + b(a + b) \\ &= aa + ab + ba + bb \\ &= a + ab + ba + b. \end{aligned}$$

$$\text{Hence } ab + ba = 0, \text{ so that by (ii), } ab = ba.$$

Note:

A ring R is called a **Boolean ring** if $a^2 = a$ for all $a \in R$. For example $(\wp(S), \Delta, \cap)$ is a Boolean ring.

Problem: 2

Complete the Cayley table for the ring $R = \{a, b, c, d\}$

+	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	b
d	d	c	b	a

.	a	b	c	d
a	a	a	a	a
b	a	b		
c	a		a	
d	a	b	c	

Solution:

First we shall compute cb .

$$cb = (b + d)b \quad (\text{from addition table})$$

$$= bb + db$$

$$= b + b \quad (\text{from multiplication table})$$

$$= a. \quad (\text{from addition table})$$

$$\text{Now, } cc = c(b + d) = cb + cd = a + a = a.$$

Space for Hints

$$bc = (b + d)c = cc + dc = a + c = c.$$

$$bd = b(b + c) = bb + bc = b + c = d.$$

$$dd = (b + c)d = bd + cd = d + a = d.$$

Hence the completed table for multiplication is

.	a	b	c	d
a	a	a	a	a
b	a	b	c	d
c	a	a	a	a
d	a	b	c	d

Exercises:

1. Given any positive integer n show that there exists a ring with n elements.
2. Let R be a ring and n any positive integer. Let $a_1, a_2, \dots, a_n \in R$. Prove that

$$(-a_1)(-a_2)\dots(-a_n) = \begin{cases} a_1 a_2 \dots a_n & \text{if } n \text{ is even} \\ -a_1 a_2 \dots a_n & \text{if } n \text{ is odd} \end{cases}$$

3. Complete the Cayley table for the ring $R = \{a, b, c, d\}$.

+	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	b
d	d	c	b	a

.	a	b	c	d
a	a	a	a	a
b	a	b		
c	a			c
d	a	b	c	

Show that in this ring $xx = x$ for all $x \in R$.

4. Prove by induction that

$$a(b_1 + b_2 + \dots + b_n) = ab_1 + ab_2 + \dots + ab_n.$$

UNIT – 6

VECTOR SPACES

6.0 Introduction:

Upto this point we have been introduced to two basic algebraic systems namely **groups** and **rings**. In this chapter we introduce another algebraic system known as **vector spaces**. The idea of a vector arises in the study of various physical applications. Many physical entities like mass, temperature etc., are characterized in terms of a real number and are called **scalars**. Other physical entries such as the velocity of a particle or force acting at a point are determined only when both magnitude and direction are specified. Such entities are called **vectors**. Since the concept of direction is geometrical a vector can be represented geometrically by a line segment whose direction is that of the given vector and whose length represents the magnitude of the vector. Two vectors $\mathbf{u}$ and $\mathbf{v}$ passing through a point $\mathbf{O}$ can be added by the usual parallelogram law of forces and we obtain the vector $\mathbf{u} + \mathbf{v}$. The vector of zero magnitude is zero vector denoted by $\mathbf{0}$ and clearly $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$. If a vector $\mathbf{u}$ is represented by a line segment $\overline{AB}$, then the vector represented by the line segment, $\overline{BA}$ is called the **negative** of $\mathbf{u}$ and is denoted by $-\mathbf{u}$ and it is clear that $\mathbf{u} + (-\mathbf{u}) = (-\mathbf{u}) + \mathbf{u} = \mathbf{0}$. Further this addition of vectors is commutative and associative. Hence the set V of vectors at a point $\mathbf{O}$ in a plane is an abelian group with respect to addition.

If $\mathbf{u}$ is a vector then $\mathbf{u} + \mathbf{u} = 2\mathbf{u}$ is evidently a vector in the same direction as $\mathbf{u}$, but of twice its magnitude. This introduces a new concept of multiplication of a vector by a scalar, the resulting product being a vector.

Thus given any real number α and a vector $\mathbf{u}$ passing through $\mathbf{O}$ then $\alpha\mathbf{u}$ is the vector whose direction is either the same as that of $\mathbf{u}$ or opposite to that of $\mathbf{u}$ according as $\alpha > 0$ or $\alpha < 0$ and whose magnitude is $|\alpha|$ times the magnitude of $\mathbf{u}$. This association gives rise to a map from $\mathbf{R} \times V$ to V given by $(\alpha, \mathbf{u}) \rightarrow \alpha\mathbf{u}$. It can be easily be verified that

$(\alpha + \beta)u = \alpha u + \beta u$ and $\alpha(u + v) = \alpha u + \alpha v$ where $u, v \in V$ and $\alpha, \beta \in R$. This ideas motivate the following abstract definition of a vector space V over a field F .

6.1 Definition and Examples:

Definition: 6.1.1

A non – empty set V is said to be a vector space over a field F if

- (i) V is an abelian group under an operation called **addition** which we denote by $+$.
- (ii) For every $\alpha \in F$ and $v \in V$, there is defined an element αv in V subject to the following conditions.
 - (a) $\alpha(u + v) = \alpha u + \alpha v$ for all $u, v \in V$ and $\alpha \in F$.
 - (b) $(\alpha + \beta)u = \alpha u + \beta u$ for all $u \in V$ and $\alpha, \beta \in F$.
 - (c) $\alpha(\beta u) = (\alpha\beta)u$ for all $u \in V$ and $\alpha, \beta \in F$.
 - (d) $1u = u$ for all $u \in V$.

Remarks:

1. The elements of F are called **scalars** and the elements of V are called **vectors**.
2. The rule which associates with each scalar $\alpha \in F$ and a vector $v \in V$, a vector αv is called the **scalar multiplication**. Thus a scalar multiplication gives rise to a function form $F \times V \rightarrow V$ defined by $(\alpha, v) \rightarrow \alpha v$.

Examples:

1. $R \times R$ is a vector space over R under addition and scalar multiplication defined by $(x_1, x_2) + (y_1, y_2) = (x_1 + y_1, x_2 + y_2)$ and $\alpha(x_1, x_2) = (\alpha x_1, \alpha x_2)$

Proof:

Clearly the binary operation $+$ is commutative and associative and $(0, 0)$ is the zero element.

The inverse of (x_1, x_2) is $(-x_1, -x_2)$. Hence $(R \times R, +)$ is an abelian group.

Now, let $u = (x_1, x_2)$ and $v = (y_1, y_2)$ and let $\alpha, \beta \in R$.

Then $\alpha(u + v) = \alpha[(x_1, x_2) + (y_1, y_2)]$

$$= \alpha(x_1 + y_1, x_2 + y_2)$$

$$\begin{aligned}
&= (\alpha x_1 + \alpha y_1, \alpha x_2 + \alpha y_2) \\
&= (\alpha x_1, \alpha x_2) + (\alpha y_1, \alpha y_2) \\
&= \alpha(x_1, x_2) + \alpha(y_1, y_2) \\
&= \alpha u + \alpha v.
\end{aligned}$$

$$\begin{aligned}
\text{Now, } (\alpha + \beta)u &= (\alpha + \beta)(x_1, x_2) \\
&= ((\alpha + \beta)x_1, (\alpha + \beta)x_2) \\
&= (\alpha x_1 + \beta x_1, \alpha x_2 + \beta x_2) \\
&= (\alpha x_1, \alpha x_2) + (\beta x_1, \beta x_2) \\
&= \alpha(x_1, x_2) + \beta(x_1, x_2) \\
&= \alpha u + \beta u
\end{aligned}$$

$$\begin{aligned}
\text{Also } \alpha(\beta u) &= \alpha(\beta(x_1, x_2)) \\
&= \alpha(\beta x_1, \beta x_2) \\
&= (\alpha \beta x_1, \alpha \beta x_2) \\
&= (\alpha \beta)(x_1, x_2) \\
&= (\alpha \beta)u
\end{aligned}$$

Obviously $1u = u$.

$\therefore \mathbf{R} \times \mathbf{R}$ is a vector space over $\mathbf{R}$.

2. $\mathbf{R}^n = \{(x_1, x_2, \dots, x_n) / x_i \in \mathbf{R}, 1 \leq i \leq n\}$. Then $\mathbf{R}^n$ is a vector space over $\mathbf{R}$ under addition and scalar multiplication defined by

$$\begin{aligned}
&(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n) \\
&\text{and } \alpha(x_1, x_2, \dots, x_n) = (\alpha x_1, \alpha x_2, \dots, \alpha x_n).
\end{aligned}$$

Proof:

Clearly the binary operation $+$ is commutative and associative. $(0, 0, \dots, 0)$ is the zero element. The inverse of

$(x_1, x_2, \dots, x_n)$ is $(-x_1, -x_2, \dots, -x_n)$. Hence $(\mathbf{R}^n, +)$ is an abelian group. Now, let $u = (x_1, x_2, \dots, x_n)$ and

$v = (y_1, y_2, \dots, y_n)$ and let $\alpha, \beta \in \mathbf{R}$.

Then $\alpha(u + v) = \alpha[(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n)]$

$$= \alpha(x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$$

$$= (\alpha x_1 + \alpha y_1, \alpha x_2 + \alpha y_2, \dots, \alpha x_n + \alpha y_n)$$

$$= (\alpha x_1, \alpha x_2, \dots, \alpha x_n) + (\alpha y_1, \alpha y_2, \dots, \alpha y_n)$$

$$= \alpha(x_1, x_2, \dots, x_n) + \alpha(y_1, y_2, \dots, y_n)$$

$$= \alpha u + \alpha v.$$

Similarly $(\alpha + \beta)u = \alpha u + \beta u$ and $\alpha(\beta u) = (\alpha\beta)u$.

Obviously $1u = u$.

$\therefore R^n$ is vector space over R .

Note:

We denote this vector space by $V_n(R)$.

Let F be any field.

Let $F^n = \{(x_1, x_2, \dots, x_n) / x_i \in F\}$.

In F^n we defined addition and scalar multiplication as in Example 2.

Then F^n is a vector space over F and we denote this vector space by $V_n(F)$.

Note:

In this example if we take $n = 1$ then we see that any field F is a vector space over itself. The addition and scalar multiplication in this vector space are simply the addition and multiplication of the field F .

C is a vector space over the field R .

Here addition is the usual addition in C and the scalar multiplication is the usual multiplication of a real number and a complex number. (ie)

$$(x_1 + ix_2) + (y_1 + iy_2) = (x_1 + y_1) + i(x_2 + y_2) \text{ and } \alpha(x_1 + ix_2) = \alpha x_1 + i\alpha x_2$$

Proof:

Clearly $(C, +)$ is an abelian group. Also the remaining axioms of a vector space are true since the scalars and vectors involved are complex numbers and further the operations are usual addition and multiplication. Hence C is a vector space over R .

Let $V = \{a + b\sqrt{2} / a, b \in Q\}$. Then V is a vector space over Q under addition and multiplication.

Proof:

Obviously V is an abelian group under usual addition.

The remaining axioms of a vector space are true since the scalars and vectors are real numbers and the operations are usual addition and multiplication. Hence V is a vector space over Q .

6. Let F be a field. Then $F[x]$, the set of all polynomials over F , is a vector space over F under the addition of polynomials and scalar multiplication defined by

$$\alpha(a_0 + a_1x + \dots + a_nx^n) = \alpha a_0 + \alpha a_1x + \dots + \alpha a_nx^n.$$

7. The set V of all polynomials of degree $\leq n$ including the zero polynomial in $F[x]$ is a vector space over the field F under the addition and scalar multiplication defined as in Example 6.

Proof:

Let $f, g \in V$. Then f and g are polynomials of degree $\leq n$.

$\therefore f + g$ and αf are of degree $\leq n$.

$\therefore f + g, \alpha f \in V$.

The other axioms of a vector space can easily be verified. Hence V is a vector space over F .

8. The set $M_2(\mathbb{R})$ of all 2×2 matrices is a vector space over $\mathbb{R}$ under matrix addition and scalar multiplication defined by

$$\alpha \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \alpha a & \alpha b \\ \alpha c & \alpha d \end{bmatrix}.$$

9. Let V be the set of all functions from $\mathbb{R}$ to $\mathbb{R}$. Let $f, g \in V$. We define $(f + g)(x) = f(x) + g(x)$ and $(\alpha f)(x) = \alpha[f(x)]$.

V is a vector space over $\mathbb{R}$. (verify)

10. Let V denote the set of all solutions of the differential equation

$$2 \frac{d^2y}{dx^2} - 7 \frac{dy}{dx} + 3y = 0.$$

Then V is a vector space over $\mathbb{R}$.

Proof:

Let $f, g \in V$ and $\alpha \in \mathbb{R}$. Then $2 \frac{d^2f}{dx^2} - 7 \frac{df}{dx} + 3f = 0$

and $2 \frac{d^2g}{dx^2} - 7 \frac{dg}{dx} + 3g = 0$.

$$\therefore 2 \left[\frac{d^2f}{dx^2} + \frac{d^2g}{dx^2} \right] - 7 \left[\frac{df}{dx} + \frac{dg}{dx} \right] + 3(f + g) = 0.$$

Hence $f + g \in V$.

$$\text{Also } 2\frac{d^2}{dx^2}(\alpha f) - 7\frac{d}{dx}(\alpha f) + 3(\alpha f) = 0.$$

Hence $\alpha f \in V$.

Since the operations are usual addition and scalar multiplication, the axioms of vector space are true.

Hence V is a vector space over $\mathbb{R}$.

11. Any sequence of real numbers $a_1, a_2, \dots, a_n, \dots$ is usually denoted by the symbol (a_n) . Let V denote the set of all sequences of real numbers. V is a vector space over the field of real numbers. The addition and scalar multiplication are defined by $(a_n) + (b_n) = (a_n + b_n)$ and $\alpha(a_n) = (\alpha a_n)$.

12. Let $V = \{0\}$. V is a vector space over any field F under the obvious operations of addition and scalar multiplication.

13. $\mathbb{R}$ is not a vector space over $\mathbb{C}$. Clearly $(\mathbb{R}, +)$ is an abelian group. But the scalar multiplication is not defined, for if $\alpha = a + ib \in \mathbb{C}$ and $u \in \mathbb{R}$, then $\alpha u = au + ibu \notin \mathbb{R}$.
 $\therefore \mathbb{R}$ is not vector space over $\mathbb{C}$.

14. Consider $\mathbb{R} \times \mathbb{R}$ with usual addition. We define scalar multiplication by $\alpha(x, y) = (\alpha x, \alpha^2 y)$. Then $\mathbb{R} \times \mathbb{R}$ is not a vector space over $\mathbb{R}$. clearly $\mathbb{R} \times \mathbb{R}$ with usual addition is an abelian group.
 $(\alpha + \beta)(x, y) = ((\alpha + \beta)x, (\alpha + \beta)^2 y) = (\alpha x + \beta x, \alpha^2 y + \beta^2 y + 2\alpha\beta y)$ Also,
 $\alpha(x, y) + \beta(x, y) = (\alpha x, \alpha^2 y) + (\beta x, \beta^2 y) = (\alpha x + \beta x, \alpha^2 y + \beta^2 y)$
Hence $(\alpha + \beta)(x, y) \neq \alpha(x, y) + \beta(x, y)$.
 $\therefore \mathbb{R} \times \mathbb{R}$ is not a vector space over $\mathbb{R}$.

15. Consider $\mathbb{R} \times \mathbb{R}$ with usual addition. Define the scalar multiplication as $\alpha(a, b) = (0, 0)$. Clearly $\mathbb{R} \times \mathbb{R}$ is an abelian group. Also,

(i) $\alpha(u + v) = 0$ and
 $\alpha u + \alpha v = 0 + 0 = 0$; so that
 $\alpha(u + v) = \alpha u + \alpha v$.

(ii) Similarly $(\alpha + \beta)u = \alpha u + \beta u = 0$.

(iii) $\alpha(\beta u) = (\alpha\beta)u = 0$.

However $1(a, b) = (0, 0)$.

Hence it is not a vector space.

Note:

In this example all the axioms except the axiom $1u = u$ are satisfied. Hence the axiom $1u = u$ cannot be derived from the other axioms of the vector space. Thus the axiom $1u = u$ is independent of the other axioms of the vector space. We say that the axiom $1u = u$ is **irredundant**.

16. Let V be the set of all ordered pairs of real numbers. Addition and multiplication are defined by

$$(x, y) + (x_1, y_1) = (x + x_1, y + y_1) \text{ and}$$

$$\alpha(x, y) = (x, \alpha y) \text{ where } x, y, x_1, y_1 \text{ and } \alpha \text{ are real numbers. Then}$$

V is not a vector space over $\mathbb{R}$.

Clearly V is an abelian group under the operation $+$ defined above.

Let $\alpha, \beta \in \mathbb{R}$ and $(x, y) \in V$.

$$\text{Now } (\alpha + \beta)(x, y) = (x, (\alpha + \beta)y)$$

$$= (x, \alpha y + \beta y)$$

$$\text{Also } \alpha(x, y) + \beta(x, y) = (x, \alpha y) + (x, \beta y)$$

$$= (2x, \alpha y + \beta y)$$

$$\therefore (\alpha + \beta)(x, y) \neq \alpha(x, y) + \beta(x, y).$$

Hence V is not a vector space over $\mathbb{R}$.

17. Let $\mathbb{R}^+$ be the set of all positive real numbers. Define addition and scalar multiplication as follows

$$u + v = uv \text{ for all } u, v \in \mathbb{R}^+; \alpha u = u^\alpha \text{ for all } u \in \mathbb{R}^+ \text{ and } \alpha \in \mathbb{R}.$$

Then $\mathbb{R}^+$ is a real vector space.

Proof:

Clearly $(\mathbb{R}^+, +)$ is an abelian group with identity. (verify).

$$\text{Now, } \alpha(u + v) = \alpha(uv)$$

$$= (uv)^\alpha$$

$$= (u^\alpha v^\alpha)$$

$$= \alpha u + \alpha v.$$

$$(\alpha + \beta)u = u^{\alpha+\beta}$$

$$= u^\alpha u^\beta$$

$$= \alpha u + \beta u.$$

$$\alpha(\beta u) = \alpha u^\beta$$

$$= (u^\beta)^\alpha$$

$$= u^{\beta\alpha}$$

$$= u^{\alpha\beta}$$

$$= (\alpha\beta)u.$$

$$\text{Also, } 1u = u1 = u.$$

$\therefore R^+$ is a vector space over R .

Exercises:

1. Prove that the following are vector spaces over the stated fields under usual operations.

(i) $V = \{a + b\sqrt{2} + c\sqrt{3} / a, b, c \in Q\}$ over Q .

(ii) $V = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} / a, b \in R \right\}$ over R .

(iii) $V =$ set of all constant functions from R to R .

(iv) $V =$ set of all expressions of the form $ax + by + cz$ where $a, b, c \in R$, under usual addition and scalar multiplication over R .

(v) Let F be any field and X any non – empty set. Then F^X is a vector space over F .

(vi) $V = \{0, 1, 2, x+1, x+2, 2x+1, 2x+2, x, 2x\} \subset Z_3[x]$ over Z_3 .

(vii) $V = \{(a, b, c) / a, b, c \in Q \text{ and } a + 2b = 3c\}$ over Q .

(viii) $V = \{f / f \in F[x] \text{ and } f(\alpha) = 0\}$ over F .

2. Show that the following are not vector spaces under usual addition and scalar multiplication.

(a) Z over Q

(b) $Q[x]$ over R

(c) Z over Z_5 .

- (d) The set of all polynomials in $F[x]$ of degree n where n is a fixed positive integer over F .
- (e) The set of all polynomials in $F[x]$ of degree $\geq n$ where n is a fixed positive integer over F .
3. Show that the following are not vector spaces.
- (a) $\mathbf{R} \times \mathbf{R}$ with addition defined by $(a, b) + (c, d) = (a - c, b - d)$ and usual scalar multiplication.
- (b) $\mathbf{R} \times \mathbf{R}$ with addition defined by $(a, b) + (c, d) = (0, b + d)$ and usual scalar multiplication.
- (c) $\mathbf{R} \times \mathbf{R}$ with usual addition and scalar multiplication defined by $\alpha(a, b) = (0, \alpha b)$.
- (d) $\mathbf{R} \times \mathbf{R}$ with usual addition and scalar multiplication defined by $\alpha(a, b) = (|\alpha|a, |\alpha|b)$.
- (e) $V = \{0, x + 4, 2x + 3, 3x + 1\} \subset Z_5[x]$ w.r.t. addition modulo 5 and scalar multiplication modulo 5.
- (f) $\mathbf{R} \times \mathbf{R}$ with addition defined by $(a, b) + (c, d) = (ac, bd)$ and usual scalar multiplication.
- (g) $\mathbf{R} \times \mathbf{R}$ with addition defined by $(a, b) + (c, d) = (a + c, b + d)$ and scalar multiplication defined by $k(a, b) = (k^2a, k^2b)$.
4. V is the set of polynomials in $R[x]$ of degree ≤ 1 . If $f = \alpha_0 + \alpha_1x$ and $g = \beta_0 + \beta_1x$, then we define $f + g = (\alpha_0 + \beta_0) + (\alpha_0\beta_1 + \alpha_1\beta_0)x$ and $\alpha f = \alpha\alpha_0 + \alpha\alpha_1x$. Show that V is not a vector space over R .
5. Let X be any set. Show that $\wp(X)$ is a vector space over Z_2 under the following operations.

$$A + B = A \Delta B, \alpha A = \begin{cases} \Phi & \text{if } \alpha = 0 \\ A & \text{if } \alpha = 1 \end{cases}$$

6. Let U and W be vector spaces over the same field F .
 $V = \{(u, v) / u \in U, v \in W\}$. Show that V is a vector space over F with
 addition and scalar multiplication defined by
 $(u, w) + (u', w') = (u + u', w + w')$ and $k(u, w) = (ku, kw)$ where
 $u, u' \in U; w, w' \in W$ and $k \in F$.

Remark:

Commutativity of addition in a vector space can be derived from the other axioms of the vector space (ie), the axiom of commutativity of addition in a vector space is redundant, for,

$$\begin{aligned}(1+1)(u+v) &= 1(u+v) + 1(u+v) \\ &= 1u + 1v + 1u + 1v \\ &= u + v + u + v\end{aligned}$$

$$\begin{aligned}\text{Also, } (1+1)(u+v) &= (1+1)u + (1+1)v \\ &= u + u + v + v.\end{aligned}$$

$$\begin{aligned}\therefore u + v + u + v &= u + u + v + v \\ \therefore v + u &= u + v.\end{aligned}$$

Theorem: 6.1

Let V be a vector space over a field F . Then

- (i) $\alpha 0 = 0$ for all $\alpha \in F$.
- (ii) $0v = 0$ for all $v \in V$.
- (iii) $(-\alpha)(v) = \alpha(-v) = -(\alpha v)$ for all $\alpha \in F$ and $v \in V$.
- (iv) $\alpha v = 0 \Rightarrow \alpha = 0$ or $v = 0$.

Proof:

$$(i) \quad \alpha 0 = \alpha(0+0) = \alpha 0 + \alpha 0. \text{ Hence } \alpha 0 = 0.$$

$$(ii) \quad 0v = (0+0)v = 0v + 0v. \text{ Hence } 0v = 0.$$

$$(iii) \quad 0 = [\alpha + (-\alpha)]v = \alpha v + (-\alpha)v.$$

$$\text{Hence } (-\alpha)v = -(\alpha v).$$

$$\text{Similarly } \alpha(-v) = -(\alpha v).$$

$$\text{Hence } (-\alpha)v = \alpha(-v) = -(\alpha v)$$

$$(iv) \text{ Let } \alpha v = 0. \text{ If } \alpha = 0, \text{ there is nothing to prove.}$$

$$\therefore \text{ Let } \alpha \neq 0. \text{ Then } \alpha^{-1} \in F.$$

Now, $v = 1v = (\alpha^{-1}\alpha)v = \alpha^{-1}(\alpha v) = \alpha^{-1}0 = 0$.

Exercises:

- Let V be any vector space over a field F . Show that
 - $\alpha(u - v) = \alpha u - \alpha v$.
 - $\alpha u = \alpha v$ and $\alpha \neq 0 \Rightarrow u = v$.
 - $\alpha u = \beta u$ and $u \neq 0 \Rightarrow \alpha = \beta$.
- Determine which of the following statements are true and which are false.
 - $\{(a, 0, 0) / a \in \mathbb{R}\}$ is a vector space over $\mathbb{R}$.
 - $\mathbb{R}[x]$ is a vector space over $\mathbb{C}$.
 - The set of all polynomials of even degree in $\mathbb{R}[x]$ is a vector space over $\mathbb{R}$.
 - The set of all solutions in $\mathbb{R} \times \mathbb{R}$ of the equation $x + 2y = 0$ is a vector space over $\mathbb{R}$.
 - The sum of two vectors is a vector.
 - The product of a scalar and a vector is a vector.

Answers:

2. (a) T (b) F (c) F (d) T
 (e) T (f) T.

6.2 Subspaces:

Definition: 6.2.1

Let V be a vector space over a field F . A non – empty subset W of V is called a **subspace** of V if W itself is a vector space over F under the operations of V .

Theorem: 6.2

Let V be a vector space over F . A non – empty subset W of V is a subspace of V iff W is closed with respect to vector addition and scalar multiplication in V .

Proof:

Let W be a subspace of V .

Then W itself is a vector space and hence W is closed with respect to vector addition and scalar multiplication.

Space for Hints

Conversely, let W be a non – empty subset of V such that $u, v \in W \Rightarrow u + v \in W$ and $u \in W$ and $\alpha \in F \Rightarrow \alpha u \in W$.

We prove that W is a subspace of V .

Since W is non – empty, there exists an element $u \in W$.

$$\therefore 0u = 0 \in W.$$

$$\text{Also } v \in W \Rightarrow (-1)v = -v \in W.$$

Thus W contains 0 and the additive inverse of each of its element.

Hence W is an additive subgroup of V .

$$\text{Also } u \in W \Rightarrow \alpha u \in W.$$

Since the elements of W are the elements of V the other axioms of a vector space are true in W .

Hence W is a subspace of V .

Theorem: 6.3

Let V be a vector space over a field F . A non – empty subset W of V is a subspace of V iff $u, v \in W$ and $\alpha, \beta \in F \Rightarrow \alpha u + \beta v \in W$.

Proof:

Let W be a subspace of V . Let $u, v \in W$ and $\alpha u + \beta v \in W$.

Conversely,

$$\text{Let } u, v \in W \text{ and } \alpha, \beta \in F \Rightarrow \alpha u + \beta v \in W.$$

$$\text{Taking } \alpha = \beta = 1, \text{ we get } u, v \in W \Rightarrow u + v \in W.$$

Taking $\beta = 0$, we get

$$\alpha \in F \text{ and } u \in W \Rightarrow \alpha u \in W.$$

Hence by theorem 6.2, W is a subspace of V .

Examples:

1. $\{0\}$ and V are subspaces of any vector space V . They are called the trivial subspaces of V .

2. $W = \{(a, 0, 0) / a \in \mathbb{R}\}$ is a subspace of $\mathbb{R}^3$, for, let

$$u = (a, 0, 0), v = (b, 0, 0) \in W \text{ and } \alpha, \beta \in \mathbb{R},$$

$$\text{Then } \alpha u + \beta v = \alpha(a, 0, 0) + \beta(b, 0, 0)$$

$$= (\alpha a + \beta b, 0, 0) \in W.$$

Hence W is a subspace of $\mathbb{R}^3$.

Note:

Space for Hints

Geometrically W consists of all points on the x - axis in the Euclidean 3 space.

3. In R^3 , $W = \{(ka, kb, kc) / k \in R\}$ is a subspace of R^3

For, if $u = (k_1a, k_1b, k_1c)$ and $v = (k_2a, k_2b, k_2c) \in W$ and $\alpha, \beta \in R$ then $\alpha u + \beta v = \alpha(k_1a, k_1b, k_1c) + \beta(k_2a, k_2b, k_2c)$
 $= ((\alpha k_1 + \beta k_2)a, (\alpha k_1 + \beta k_2)b, (\alpha k_1 + \beta k_2)c) \in W$

Hence W is a subspace of R^3 .

Note:

Geometrically W consists of all points of line $\frac{x}{a} = \frac{y}{b} = \frac{z}{c}$

provided a, b, c are not all zero. Thus the set of all points on a line through the origin is a subspace of R^3 . However a line not passing through the origin is not a subspace of R^3 , since the additive identity $(0, 0, 0)$ does not lie on the line.

4. $W = \{(a, b, 0) / a, b \in R\}$ is a subspace of R^3 . W consists of all points of the xy -plane.

5. Let W be the set of all points in R^3 satisfying the equation $lx + my + nz = 0$. W is a subspace of R^3 . For, let $u = (a_1, b_1, c_1)$ and $v = (a_2, b_2, c_2) \in W$ and $\alpha, \beta \in R$. Then we have

$$\ell a_1 + m b_1 + n c_1 = 0 = \ell a_2 + m b_2 + n c_2.$$

$$\text{Hence } \alpha(\ell a_1 + m b_1 + n c_1) + \beta(\ell a_2 + m b_2 + n c_2) = 0.$$

$$\text{(i.e.) } \ell(\alpha a_1 + \beta a_2) + m(\alpha b_1 + \beta b_2) + n(\alpha c_1 + \beta c_2) = 0.$$

$$\text{(i.e.) } \alpha u + \beta v \in W \text{ so that } W \text{ is a subspace of } R^3.$$

Note:

Geometrically W consists of all points on the plane $\ell x + my + nz = 0$ which passes through the origin.

6. Let $W = \{f / f \in F[x] \text{ and } f(a) = 0\}$.

(i.e) W is the set of all polynomials in $F[x]$ having a as a root where $a \in F$. Then W is a vector space over F . We observe that $x - a \in W$ and hence W is non - empty.

Let $f, g \in F[x]$ and $\alpha, \beta \in F$.

To prove that $\alpha f + \beta g \in W$ we have to show that is a root of $\alpha f + \beta g$.

Now, $(\alpha f + \beta g)(a) = \alpha f(a) + \beta g(a) = \alpha 0 + \beta 0 = 0$.

Hence a is a root of $\alpha f + \beta g$.

$\therefore \alpha f + \beta g \in W$ and hence W is a subspace of $F[x]$.

$$7. W = \left\{ \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} / a, b \in \mathbb{R} \right\} \text{ is a subspace of } M_2(\mathbb{R})$$

Problems:

Problem: 1

Prove that the intersection of two subspaces of a vector space is a subspace.

Solution:

Let A and B be two subspaces of a vector space V over a field F .

We claim that $A \cap B$ is a subspace of V .

Clearly $0 \in A \cap B$ and hence $A \cap B$ is non-empty.

Now, let $u, v \in A \cap B$ and $\alpha, \beta \in F$.

Then $u, v \in A$ and $u, v \in B$.

$\therefore \alpha u + \beta v \in A$ and $\alpha u + \beta v \in B$

(since A and B are subspaces)

$\therefore \alpha u + \beta v \in A \cap B$.

Hence $A \cap B$ is a subspace of V .

Problem: 2

Prove that the union of two subspaces of a vector space need not be a subspace.

Solution:

Let $A = \{(a, 0, 0) / a \in \mathbb{R}\}$

$B = \{(0, b, 0) / b \in \mathbb{R}\}$.

Clearly A and B are subspaces of $\mathbb{R}^3$ (example 2 of 6.3)

However $A \cup B$ is not a subspace of $\mathbb{R}^3$.

For, $(1, 0, 0)$ and $(0, 1, 0) \in A \cup B$.

But $(1, 0, 0) + (0, 1, 0) = (1, 1, 0) \notin A \cup B$.

Problem: 3

Prove that the union of two subspaces of a vector space is a subspace iff one is contained in the other.

Solution:

For proof refer theorem 2.11.

Problem: 4

If A and B are subspaces of V prove that $A + B = \{v \in V / v = a + b, a \in A, b \in B\}$ is a subspace of V . Further show that $A + B$ is the smallest subspace containing A and B . (ie) If W is any subspace of V containing A and B then W contains $A + B$.

Solution:

Let $v_1, v_2 \in A + B$ and $\alpha \in F$.

Then $v_1 = a_1 + b_1, v_2 = a_2 + b_2$ where $a_1, a_2 \in A$ and $b_1, b_2 \in B$.

Now, $v_1 + v_2 = (a_1 + b_1) + (a_2 + b_2)$
 $= (a_1 + a_2) + (b_1 + b_2) \in A + B$.

Also $\alpha(a_1 + b_1) = \alpha a_1 + \alpha b_1 \in A + B$.

Hence $A + B$ is a subspace of V . Clearly $A \subseteq A + B$ and $B \subseteq A + B$.

Now, let W be any subspace of V containing A and B .

We shall prove that $A + B \subseteq W$.

Let $v \in A + B$. Then $v = a + b$ where $a \in A$ and $b \in B$.

Since $A \subseteq W$, $a \in W$. Similarly $b \in W$.

$\therefore a + b = v \in W$.

$\therefore A + B \subseteq W$ so that $A + B$ is the smallest subspace of V containing A and B .

Problem: 5

Let A and B be subspace of a vector space V . Then $A \cap B = \{0\}$ iff every vector $v \in A + B$ can be uniquely expressed in the form $v = a + b$ where $a \in A$ and $b \in B$.

Solution:

Let $A \cap B = \{0\}$. Let $v \in A + B$.

Let $v = a_1 + b_1 = a_2 + b_2$ where $a_1, a_2 \in A$ and $b_1, b_2 \in B$.

Then $a_1 - a_2 = b_2 - b_1$.

But $a_1 - a_2 \in A$ and $b_2 - b_1 \in B$.

Hence $a_1 - a_2, b_2 - b_1 \in A \cap B$.

Since $A \cap B = \{0\}$, $a_1 - a_2 = 0$ and $b_2 - b_1 = 0$ so that $a_1 = a_2$ and $b_1 = b_2$. Hence the expression of v in the form $a + b$ where $a \in A$ and $b \in B$ is unique.

Conversely suppose that any element in $A + B$ can be uniquely expressed in the form $a + b$ where $a \in A$ and $b \in B$.

We claim that $A \cap B = \{0\}$.

If $A \cap B \neq \{0\}$, let $v \in A \cap B$ and $v \neq 0$.

Then $0 = v - v = 0 + 0$. Thus 0 has been expressed in the form $a + b$ in two different ways which is a contradiction. Hence $A \cap B = \{0\}$.

Definition: 6.2.2

Let A and B be subspace of a vector space V . Then V is called the **direct sum** of A and B if (i) $A + B = V$ (ii) $A \cap B = \{0\}$.

If V is the direct sum of A and B we write $V = A \oplus B$.

Note:

$V = A \oplus B$ iff every element of V can be uniquely expressed in the form $a + b$ where $a \in A$ and $b \in B$.

Examples:

1. In $V_3(\mathbb{R})$ let $A = \{(a, b, 0) / a, b \in \mathbb{R}\}$ and $B = \{(0, 0, c) / c \in \mathbb{R}\}$. Clearly A and B are subspace of V and $A \cap B = \{0\}$. Also let $v = (a, b, c) \in V_3(\mathbb{R})$.

Then $v = (a, b, 0) + (0, 0, c)$ so that $A + B = V_3(\mathbb{R})$.

Hence $V_3(\mathbb{R}) = A \oplus B$.

2. In $M_2(\mathbb{R})$, let A be the set of all matrices of the form $\begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix}$ and B

be the set of all matrices of the form $\begin{pmatrix} 0 & 0 \\ c & d \end{pmatrix}$. Clearly A and B are

subspaces of $M_2(\mathbb{R})$ and $A \cap B = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ and $A + B = M_2(\mathbb{R})$.

Hence $M_2(R) = A \oplus B$.

Exercises:

1. Show that the following subsets of R^3 are subspaces. Interpret them geometrically.
 - (a) $\{(a, 0, c) / a, c \in R\}$
 - (b) $\{(a, b, c) / a = b = c\}$
 - (c) $\{(a, b, c) / a = b + c\}$
 - (d) $\{(a, b, a + b) / a, b \in R\}$
2. Show that the set of all polynomials in $R[x]$ having at least one rational root is not a subspace of $R[x]$.
3. Let V be the set of all sequences in R . V is a vector space over R . Let W be the set of all convergent sequences in V and U be the set of all sequences converging to zero. Prove that (a) U is a subspace of W (b) U is a subspace of V . (c) W is a subspace of V .
4. Let W be a subspace of V and U be a subspace of W . Then show that U is a subspace of V .
5. Show that each of the following subsets of $V_3(R)$ is not a subspace.
 - (a) $S = \{(x, y, z) / x^2 + y^2 + z^2 \leq 1\}$.
 - (b) $S = \{(x, y, z) / x + y + z = 1\}$.
 - (c) $S = \{(x, y, z) / x \geq y \geq z\}$.

Theorem: 6.4

Let V be a vector space over F and W a subspace of V . Let $V/W = \{W + v / v \in V\}$. Then V/W is a vector space over F under the following operations.

- i. $(W + v_1) + (W + v_2) = W + v_1 + v_2$.
- ii. $\alpha(W + v_1) = W + \alpha v_1$.

Proof:

Since W is a subspace of V it is a subgroup of $(V, +)$.

Since $(V, +)$ is abelian, W is a normal subgroup of $(V, +)$ so that

Space for Hints

(i) is a well defined operation.

Now we shall prove that (ii) is a well defined operation.

$$W + v_1 = W + v_2 \Rightarrow v_1 - v_2 \in W$$

$$\Rightarrow \alpha(v_1 - v_2) \in W \quad (\text{since } W \text{ is a subspace})$$

$$\Rightarrow \alpha v_1 - \alpha v_2 \in W$$

$$\Rightarrow \alpha v_1 \in W + \alpha v_2$$

$$\Rightarrow W + \alpha v_1 = W + \alpha v_2$$

Hence (ii) is a well defined operation.

Now, let $W + v_1, W + v_2, W + v_3 \in V/W$

$$\text{Then } (W + v_1) + [(W + v_2) + (W + v_3)] = (W + v_1) + (W + v_2 + v_3)$$

$$= W + v_1 + v_2 + v_3$$

$$= (W + v_1 + v_2) + (W + v_3)$$

$$= [(W + v_1) + (W + v_2)] + (W + v_3)$$

Hence $+$ is associative.

$W + 0 = W \in V/W$ is the additive identity element.

For $(W + v_1) + (W + 0) = W + v_1 = (W + 0) + (W + v_1)$ for all $v_1 \in V$

Also $W - v_1$ is the additive inverse of $W + v_1$.

Hence V/W is a group under $+$.

Further $(W + v_1) + (W + v_2) = W + v_1 + v_2$

$$= W + v_2 + v_1$$

$$= (W + v_2) + (W + v_1)$$

Hence V/W is an abelian group.

Now, let $\alpha, \beta \in F$.

$$\alpha[(W + v_1) + (W + v_2)] = \alpha(W + v_1 + v_2)$$

$$= W + \alpha(v_1 + v_2)$$

$$= W + \alpha v_1 + \alpha v_2$$

$$= (W + \alpha v_1) + (W + \alpha v_2)$$

$$= \alpha(W + v_1) + \alpha(W + v_2)$$

$$(\alpha + \beta)(W + v_1) = W + (\alpha + \beta)v_1$$

$$= W + \alpha v_1 + \beta v_1$$

$$= (W + \alpha v_1) + (W + \beta v_1)$$

$$= \alpha(W + v_1) + \beta(W + v_1)$$

$$\alpha[\beta(W + v_1)] = \alpha(W + \beta v_1)$$

$$= W + \alpha\beta v_1$$

$$= (\alpha\beta)(W + v_1)$$

$$1(W + v_1) = W + 1v_1$$

$$= W + v_1$$

Hence V/W is a vector space.

The vector space V/W is called the **quotient space** of V by W .

6.3 Linear Transformation:

Definition: 6.3.1

Let V and W be vector spaces over a field F . A mapping $T: V \rightarrow W$ is called a **homomorphism** if

- (a) $T(u + v) = T(u) + T(v)$ and
- (b) $T(\alpha u) = \alpha T(u)$ where $\alpha \in F$ and $u, v \in V$.

A homomorphism T of vector spaces is also called a **linear transformation**.

- (i). If T is 1-1 then T is called **monomorphism**.
- (ii). If T is onto then T is called an **epimorphism**.
- (iii). If T is 1-1 and onto T is called an **isomorphism**.
- (iv). Two vector spaces V and W are said to be isomorphic if there exists an isomorphism T from V to W and we write $V \cong W$.
- (v). A linear transformation $T: V \rightarrow F$ is called a **linear functional**.

Examples:

1. $T: V \rightarrow W$ defined by $T(v) = 0$ for all $v \in V$ is a **trivial linear transformation**.
2. $T: V \rightarrow V$ defined by $T(v) = v$ for all $v \in V$ is the **identity linear transformation**.
3. Let V be a vector space over a field F and W a subspace of V . Then $T: V \rightarrow V/W$ defined by $T(v) = W + v$ is a linear transformation, for, $T(v_1 + v_2) = W + (v_1 + v_2)$

$$= (W + v_1) + (W + v_2)$$

$$= T(v_1) + T(v_2)$$

$$\text{Also } T(\alpha v_1) = W + \alpha v_1 = \alpha(W + v_1) = \alpha T(v_1).$$

This is called the **natural homomorphism** from V to V/W . Clearly T is onto and hence T is an epimorphism.

4. $T: V_3(\mathbb{R}) \rightarrow V_3(\mathbb{R})$ defined by $T(a, b, c) = (a, 0, 0)$ is a linear transformation.

5. Let V be the set of all polynomials of degree $\leq n$ in $\mathbb{R}[x]$ including the zero polynomial. $T: V \rightarrow V$ defined by $T(f) = \frac{df}{dx}$ is a linear transformation.

$$\begin{aligned} \text{For, } T(f+g) &= \frac{d(f+g)}{dx} = \frac{df}{dx} + \frac{dg}{dx} \\ &= T(f) + T(g). \end{aligned}$$

$$\text{Also } T(\alpha f) = \frac{d(\alpha f)}{dx} = \alpha \frac{df}{dx} = \alpha T(f).$$

6. Let V be as in Example 5.

Then $T: V \rightarrow V_{n+1}(\mathbb{R})$ defined by

$T(a_0 + a_1x + \dots + a_nx^n) = (a_0, a_1, \dots, a_n)$ is a linear transformation. For, let $f = a_0 + a_1x + \dots + a_nx^n$ and $g = b_0 + b_1x + \dots + b_nx^n$.

$$\text{Then } f+g = (a_0 + b_0) + (a_1 + b_1)x + \dots + (a_n + b_n)x^n$$

$$\begin{aligned} \therefore T(f+g) &= ((a_0 + b_0), (a_1 + b_1), \dots, (a_n + b_n)) \\ &= (a_0, a_1, \dots, a_n) + (b_0, b_1, \dots, b_n) \\ &= T(f) + T(g). \end{aligned}$$

$$\begin{aligned} \text{Also } T(\alpha f) &= (\alpha a_0, \alpha a_1, \dots, \alpha a_n) \\ &= \alpha(a_0, a_1, \dots, a_n) \\ &= \alpha T(f). \end{aligned}$$

Clearly T is 1-1 and onto and hence T is an isomorphism.

7. Let V denote the set of all sequences in $\mathbb{R}$. $T: V \rightarrow V$ defined by

$T(a_1, a_2, \dots, a_n, \dots) = (0, a_1, a_2, \dots, a_n, \dots)$ is a linear transformation.

8. $T: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defined by $T(a, b) = (2a - 3b, a + 4b)$ is a linear transformation.

Let $u = (a, b)$ and $v = (c, d)$ and $\alpha \in \mathbb{R}$.

$$\begin{aligned} \therefore T(u + v) &= T((a, b) + (c, d)) \\ &= T(a + c, b + d) \\ &= (2(a + c) - 3(b + d), (a + c) + 4(b + d)) \\ &= (2a + 2c - 3b - 3d, a + c + 4b + 4d) \\ &= (2a - 3b + 2c - 3d, a + 4b + c + 4d) \\ &= (2a - 3b, a + 4b) + (2c - 3d, c + 4d) \\ &= T(a, b) + T(c, d) \\ &= T(u) + T(v). \end{aligned}$$

Also, $T(\alpha u) = T(\alpha(a, b))$

$$\begin{aligned} &= T(\alpha a, \alpha b) \\ &= (2\alpha a - 3\alpha b, \alpha a + 4\alpha b) \\ &= \alpha(2a - 3b, a + 4b) \\ &= \alpha T(a, b) \\ &= \alpha T(u). \end{aligned}$$

$\therefore T$ is a linear transformation.

Theorem: 6.5

Let $T: V \rightarrow W$ be a linear transformation.

Then $T(V) = \{T(v) / v \in V\}$ is a subspace of W .

Proof:

Let w_1 and $w_2 \in T(V)$ and $\alpha \in F$. Then there exist $v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$.

$$\begin{aligned} \text{Hence } w_1 + w_2 &= T(v_1) + T(v_2) \\ &= T(v_1 + v_2) \in T(V). \end{aligned}$$

Similarly $\alpha w_1 = \alpha T(v_1) = T(\alpha v_1) \in T(V)$.

Hence $T(V)$ is subspace of W .

Definition: 6.3.2

Let V and W be vector space over a field F and $T: V \rightarrow W$ be a linear transformation. Then the **kernel** of T is defined to be

$\{v / v \in V \text{ and } T(v) = 0\}$ and is denoted by **ker** T .

Thus $\text{ker } T = \{v / v \in V \text{ and } T(v) = 0\}$.

For example, in example 1, $\text{ker } T = V$.

In example 2, $\text{ker } T = \{0\}$.

In example 5, $\text{ker } T$ is the set of all constant polynomials.

Note:

Let $T: V \rightarrow W$ be a linear transformation. Then T is a monomorphism iff $\text{ker } T = \{0\}$.

6.4 Fundamental Theorem of Homomorphism

Theorem: 6.6 (Fundamental theorem of homomorphism)

Statement:

Let V and W be vector spaces over a field F and $T: V \rightarrow W$ be an epimorphism.

Then (i) $\text{ker } T = V_1$ is a subspace of V and (ii) $\frac{V}{V_1} \cong W$.

Proof:

(i) Given $V_1 = \text{ker } T = \{v / v \in V \text{ and } T(v) = 0\}$.

Clearly $T(0) = 0$. Hence $0 \in \text{ker } T = V_1$.

$\therefore V_1$ is nonempty subset of V .

Let $u, v \in \text{ker } T$ and $\alpha, \beta \in F$.

$\therefore T(u) = 0$ and $T(v) = 0$.

Now $T(\alpha u + \beta v) = T(\alpha u) + T(\beta v)$

$$= \alpha T(u) + \beta T(v)$$

$$= \alpha 0 + \beta 0$$

$$= 0$$

$\therefore \alpha u + \beta v \in \text{ker } T$

$\therefore$ By theorem 6.3, $\text{ker } T$ is a subspace of V .

(ii) We define a map $\phi: \frac{V}{V_1} \rightarrow W$ by

$$\phi(V_1 + v) = T(v).$$

ϕ is well defined. Let $V_1 + v = V_1 + w$.

$$\therefore v \in V_1 + w.$$

$$\therefore v = v_1 + w \text{ where } v_1 \in V_1.$$

$$\therefore T(v) = T(v_1 + w) = T(v_1) + T(w) = 0 + T(w) = T(w)$$

$$\therefore \phi(V_1 + v) = \phi(V_1 + w)$$

ϕ is 1-1.

$$\phi(V_1 + v) = \phi(V_1 + w)$$

$$\Rightarrow T(v) = T(w)$$

$$\Rightarrow T(v) - T(w) = 0$$

$$\Rightarrow T(v) + T(-w) = 0$$

$$\Rightarrow T(v - w) = 0$$

$$\Rightarrow v - w \in \ker T = V_1$$

$$\Rightarrow v \in V_1 + w$$

$$\Rightarrow V_1 + v = V_1 + w.$$

ϕ is onto.

Let $w \in W$. Since T is onto there exists $v \in V$ such that $T(v) = w$.

$$\therefore \phi(V_1 + v) = w.$$

$$\therefore \phi(V_1 + v) = w.$$

ϕ is a homomorphism.

$$\phi[(V_1 + v) + (V_1 + w)] = \phi[V_1 + (v + w)]$$

$$= T(v + w)$$

$$= T(v) + T(w)$$

$$= \phi(V_1 + v) + \phi(V_1 + w)$$

$$\text{Also } \phi[\alpha(V_1 + v)] = \phi(V_1 + \alpha v)$$

$$= T(\alpha v)$$

$$= \alpha T(v)$$

$$= \alpha T(V_1 + v).$$

Space for Hints

$\therefore \varphi$ is an isomorphism from $\frac{V}{V_1}$ onto W .

$\therefore \frac{V}{V_1} \cong W$.

Theorem: 6.7

Let V be a vector space over a field F .

Let A and B be subspaces of V .

Then $\frac{A+B}{A} \cong \frac{B}{A \cap B}$.

Proof:

We know that $A+B$ is a subspace of V containing A .

Hence $\frac{A+B}{A}$ is also a vector space over F .

An element of $\frac{A+B}{A}$ is of the form $A+(a+b)$ where $a \in A$ and $b \in B$. But $A+a = A$.

Hence an element of $\frac{A+B}{A}$ is of the form $A+b$.

Now, consider $f: B \rightarrow \frac{A+B}{A}$ defined by $f(b) = A+b$.

Clearly f is onto.

$$\begin{aligned}\text{Also } f(b_1 + b_2) &= A + (b_1 + b_2) \\ &= (A + b_1) + (A + b_2) \\ &= f(b_1) + f(b_2)\end{aligned}$$

$$\begin{aligned}\text{and } f(\alpha b_1) &= A + \alpha b_1 \\ &= \alpha(A + b_1) \\ &= \alpha f(b_1)\end{aligned}$$

Hence f is a linear transformation.

Let K be the kernel of f .

Then $K = \{b/b \in B, A+b = A\}$.

Now, $A+b = A$ iff $b \in A$. Hence $K = A \cap B$.

Hence by theorem 6.6, $\frac{B}{A \cap B} \cong \frac{A+B}{A}$.

Theorem: 6.8

Let V and W be vector space over a field F . Let $L(V, W)$ represent the set of all linear transformations from V to W . Then $L(V, W)$ itself is a vector space over F under addition and scalar multiplication defined by $(f + g)(v) = f(v) + g(v)$ and $(\alpha f)(v) = \alpha f(v)$.

Proof:

Let $f, g \in L(V, W)$ and $v_1, v_2 \in V$.

$$\text{Then } (f + g)(v_1 + v_2) = f(v_1 + v_2) + g(v_1 + v_2)$$

$$= f(v_1) + f(v_2) + g(v_1) + g(v_2)$$

$$= f(v_1) + g(v_1) + f(v_2) + g(v_2)$$

$$= (f + g)(v_1) + (f + g)(v_2)$$

$$\text{Also } (f + g)(\alpha v) = f(\alpha v) + g(\alpha v)$$

$$= \alpha f(v) + \alpha g(v)$$

$$= \alpha[f(v) + g(v)]$$

$$= \alpha(f + g)(v).$$

Hence $(f + g) \in L(V, W)$.

$$\text{Now, } (\alpha f)(v_1 + v_2) = (\alpha f)(v_1) + (\alpha f)(v_2)$$

$$= \alpha f(v_1) + \alpha f(v_2)$$

$$= \alpha[f(v_1) + f(v_2)]$$

$$= \alpha f(v_1 + v_2).$$

$$\text{Also } (\alpha f)(\beta v) = \alpha[f(\beta v)] = \alpha[\beta f(v)]$$

$$= \beta[\alpha f(v)] = \beta[(\alpha f)(v)].$$

Hence $\alpha f \in L(V, W)$.

Addition defined on $L(V, W)$ is obviously commutative and associative.

The function $f: V \rightarrow W$ defined by $f(v) = 0$ for all $v \in V$ is clearly a linear transformation and is the additive identity of $L(V, W)$.

Further $(-f): V \rightarrow W$ defined by

$$(-f)(v) = -f(v) \text{ is the additive inverse of } f.$$

Thus $L(V, W)$ is an abelian group under addition.

The remaining axioms for a vector space can be easily verified.

Hence $L(V, W)$ is a vector space over F .

Exercises:

1. Let V and W be vector spaces over a field F . Show that a mapping $T: V \rightarrow W$ is a linear transformation iff $T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$ for all $v_1, v_2 \in V$ and $\alpha, \beta \in F$.
2. For each of the linear transformations T given in examples of section 6.3 determine where it is one – one and onto. Also find its kernel.
3. Find the kernel of the following linear transformations.
 - (i) $T: V_4(\mathbb{R}) \rightarrow V_4(\mathbb{R})$ defined by $T(x_1, x_2, x_3, x_4) = (x_1, 0, x_3, 0)$.
 - (ii) $T: V_3(\mathbb{R}) \rightarrow V_3(\mathbb{R})$ defined by $T(a, b, c) = (a, b, 0)$.
4. Let V and W be two vector spaces over a field F and $T: V \rightarrow W$ be a linear transformation. Show that if V_1 is a subspace of V , $T(V_1)$ is a subspace of W and if W_1 is a subspace of W , $T^{-1}(W_1)$ is a subspace of V .
5. Prove that $T: V_3(\mathbb{R}) \rightarrow \mathbb{R}$ defined by $T(x, y, z) = x^2 + y^2 + z^2$ is not a linear transformation.
6. Determine which of the following statements are true.
 - (a) Any line passing through the origin is a subspace of $V_3(\mathbb{R})$.
 - (b) Any plane passing through the origin is a subspace of $V_3(\mathbb{R})$.
 - (c) If A and B are subspace of V then
 - (i) $A + B$ is a subspace of V .
 - (ii) A is a subspace of $A + B$.
 - (iii) B is a subspace of $A + B$.
 - (iv) Every element of $A + B$ can be uniquely written in the form $a + b$ where $a \in A$ and $b \in B$.
 - (v) (iv) is true iff $A \cap B = \{0\}$

Answers:

2. 1. Not 1 – 1, not onto; kernel V .
2. 1 – 1 and onto; kernel $= \{0\}$.

3. Not 1-1, onto kernel = W.

4. Not 1-1, not onto;

Kernel = $\{(0, b, c) / b, c \in \mathbb{R}\}$

5. Not 1-1 not onto; kernel = all constant polynomials

6. 1-1 and onto; kernel = $\{0\}$

7. 1-1, not onto;

Kernel = $\{0\}$

3. (i) Not 1-1, not onto;

Kernel $\{(0, a, 0, b) / a, b \in \mathbb{R}\}$

(ii) Not 1-1, not onto;

kernel $\{(0, 0, c) / c \in \mathbb{R}\}$.

6. (a) T

(b) T

(c) [i] T

[ii] T

[iii]

T

[iv] F

[v] T.

UNIT – 7

SPAN OF A SET

Definition: 7.1.1

Let V be a vector space over a field F .

Let $v_1, v_2, \dots, v_n \in V$. Then an element of the form $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$ where $\alpha_i \in F$ is called a **linear combination** of the vectors $v_1, v_2, \dots, v_n$.

Definition: 7.1.2

Let S be a non – empty subset of a vector space V . Then the set of all linear combinations of finite sets of elements of S is called the **linear span** of S and is denoted by $L(S)$.

Note:

Any element of $L(S)$ is of the form $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$ where $\alpha_1, \alpha_2, \dots, \alpha_n \in F$.

Theorem: 7.1

Let V be a vector space over a field F and S be a non-empty subset of V . Then

- i. $L(S)$ is a subspace of V .
- ii. $S \subseteq L(S)$.
- iii. If W is any subspace of V such that $S \subseteq W$, then $L(S) \subseteq W$ (ie), $L(S)$ is the smallest subspace of V containing S .

Proof:

(i) Let $v, w \in L(S)$ and $\alpha, \beta \in F$.

Then $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$ where $v_i \in S$ and $\alpha_i \in F$.

Also, $w = \beta_1 w_1 + \beta_2 w_2 + \dots + \beta_m w_m$ where $w_j \in S$ and $\beta_j \in F$. Now,

$$\begin{aligned} \alpha v + \beta w &= \alpha(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n) + \beta(\beta_1 w_1 + \beta_2 w_2 + \dots + \beta_m w_m) \\ &= (\alpha\alpha_1)v_1 + \dots + (\alpha\alpha_n)v_n + (\beta\beta_1)w_1 + \dots + (\beta\beta_m)w_m. \end{aligned}$$

$\therefore \alpha v + \beta w$ is also a linear combination of a finite number of elements of S .

Hence $\alpha v + \beta w \in L(S)$.

$\therefore L(S)$ is a subspace of V .

- (ii) Let $u \in S$.
Then $u = 1u \in L(S)$. Hence $S \subseteq L(S)$.
- (iii) Let W be any subspace of V such that $S \subseteq W$. We claim that $L(S) \subseteq W$.
Let $u \in L(S)$
Then $u = \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n$ where $u_i \in S$ and $\alpha_i \in F$.
Since $S \subseteq W$ we have $u_1, u_2, \dots, u_n \in W$.
 $\therefore \alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n \in W$. (since W is a subspace of V .)
 $\therefore u \in W$. Hence $L(S) \subseteq W$.

Note:

$L(S)$ is called the subspace **spanned** (generated) by the set S .

Examples:

1. In $V_3(\mathbb{R})$ let $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$ and $e_3 = (0, 0, 1)$

(a) Let $S = \{e_1, e_2\}$

$$\begin{aligned} \text{Then } L(S) &= \{\alpha e_1 + \beta e_2 / \alpha, \beta \in \mathbb{R}\} \\ &= \{(\alpha, \beta, 0) / \alpha, \beta \in \mathbb{R}\} \end{aligned}$$

(b) Let $S = \{e_1, e_2, e_3\}$. Then $L(S) = \{\alpha e_1 + \beta e_2 + \gamma e_3 / \alpha, \beta, \gamma \in \mathbb{R}\}$
 $= \{(\alpha, \beta, \gamma) / \alpha, \beta, \gamma \in \mathbb{R}\}$
 $= V_3(\mathbb{R})$

Thus $V_3(\mathbb{R})$ is spanned by $\{e_1, e_2, e_3\}$.

2. In $V_n(\mathbb{R})$, let $e_1 = (1, 0, \dots, 0)$,

$$e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, 0, \dots, 1).$$

Let $S = \{e_1, e_2, \dots, e_n\}$. Then

$$\begin{aligned} L(S) &= \{\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n / \alpha_i \in \mathbb{R}\} \\ &= \{(\alpha_1, \alpha_2, \dots, \alpha_n) / \alpha_i \in \mathbb{R}\} \\ &= V_n(\mathbb{R}) \end{aligned}$$

Thus $V_n(\mathbb{R})$ is spanned by $\{e_1, e_2, \dots, e_n\}$.

Theorem: 7.2

Let V be a vector space over a field F . Let $S, T \subseteq V$, then

Space for Hints

- (a) $S \subseteq T \Rightarrow L(S) \subseteq L(T)$.
- (b) $L(S \cup T) = L(S) + L(T)$.
- (c) $L(S) = S$ iff S is a subspace of V .

Proof:

(a) Let $S \subseteq T$. Let $s \in L(S)$. Then $s = \alpha_1 s_1 + \alpha_2 s_2 + \dots + \alpha_n s_n$ where $s_i \in S$ and $\alpha_i \in F$.

Now, since $S \subseteq T$, $s_i \in T$.

Hence $\alpha_1 s_1 + \alpha_2 s_2 + \dots + \alpha_n s_n \in L(T)$.

Thus $L(S) \subseteq L(T)$.

(b) Let $s \in L(S \cup T)$.

Then $s = \alpha_1 s_1 + \alpha_2 s_2 + \dots + \alpha_n s_n$ where $s_i \in S \cup T$ and $\alpha_i \in F$.

Without loss of generality we can assume that

$$s_1, s_2, \dots, s_m \in S$$

and

$$s_{m+1}, \dots, s_n \in T.$$

Hence $\alpha_1 s_1 + \alpha_2 s_2 + \dots + \alpha_m s_m \in L(S)$ and

$$\alpha_{m+1} s_{m+1} + \dots + \alpha_n s_n \in L(T)$$

$$\therefore s = (\alpha_1 s_1 + \dots + \alpha_m s_m) + (\alpha_{m+1} s_{m+1} + \dots + \alpha_n s_n) \in L(S) + L(T)$$

Hence $L(S \cup T) \subseteq L(S) + L(T)$.

Also by (a) $L(S) \subseteq L(S \cup T)$ and $L(T) \subseteq L(S \cup T)$.

Hence $L(S) + L(T) \subseteq L(S \cup T)$

Hence $L(S \cup T) = L(S) + L(T)$.

(c) Let $L(S) = S$. By theorem 7.1 $L(S) = S$ is a subspace of V .

Conversely, let S be a subspace V . Then the smallest subspace containing S is S itself. Hence $L(S) = S$.

Corollary:

$$L[L(S)] = L(S).$$

Exercises:

1. Find $L(S)$ in the following cases.

a) $S = \{(1,0), (0,1)\}$ in $V_2(R)$

b) $S = \{(1,0,0), (2,0,0), (3,0,0)\}$ in $V_3(\mathbb{R})$.

c) $S = \{(1,2,3), (2,3,1), (3,1,2)\}$ in $V_3(\mathbb{R})$.

d) $S = \{(2,0)\}$ in $V_2(\mathbb{R})$.

e) $S = \{1, x, x^2, \dots, x^n\}$ in $\mathbb{R}[x]$.

f) $S = \{(1,2,3)\}$ in $\mathbb{Z}_5 \times \mathbb{Z}_5 \times \mathbb{Z}_5$ over $\mathbb{Z}_5$.

g) $S = \{(0,1,2), (1,2,0)\}$ in $\mathbb{Z}_3 \times \mathbb{Z}_3 \times \mathbb{Z}_3$ over $\mathbb{Z}_3$.

h) $S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \right\}$ in $M_2(\mathbb{R})$.

2. Let V be a vector space over a field F and

$$S = \{v_1, v_2, \dots, v_n\} \subseteq V.$$

$$\text{Let } S_1 = \{\alpha_1 v_1, \alpha_2 v_2, \dots, \alpha_n v_n\} \text{ where } \alpha_i \in F - \{0\}.$$

$$\text{Let } S_2 = \{v_1 + \alpha v_2, v_2, v_3, \dots, v_n\} \text{ where } \alpha \in F.$$

Show that $L(S) = L(S_1) = L(S_2)$.

3. Show that in $V_2(\mathbb{R})$, $(3,7) \in L(\{(1,2), (0,1)\})$

Answers:

1. (a) $V_2(\mathbb{R})$ (b) $\{(x,0,0)/x \in \mathbb{R}\}$

(c) $V_3(\mathbb{R})$ (d) $\{(x,0)/x \in \mathbb{R}\}$

(e) The set of all polynomials of degree $\leq n$ and zero polynomial.

(f) $\{(0,0,0), (1,2,3), (2,4,1), (3,1,4), (4,3,2)\}$

(g) $\{(0,0,0), (0,1,2), (1,2,0), (1,0,2), (0,2,1), (2,1,0), (2,0,1), (1,1,1), (2,2,2)\}$

(h) $\left\{ \begin{pmatrix} a & 0 \\ b & 0 \end{pmatrix} \mid a, b \in \mathbb{R} \right\}$

7.2 Linear Independence:

In $V_3(\mathbb{R})$, let $S = \{e_1, e_2, e_3\}$. We have seen that $L(S) = V_3(\mathbb{R})$.

Thus S is a subset of $V_3(\mathbb{R})$ which spans the whole space $V_3(\mathbb{R})$.

Definition: 7.2.1

Let V be a vector space over a field F . V is said to be **finite dimensional** if there exists a finite subset S of V such that $L(S) = V$.

Examples:

1. $V_3(\mathbb{R})$ is a finite – dimensional vector space.

2. $V_n(R)$ is a finite dimensional vector space, since $S = \{e_1, e_2, \dots, e_n\}$ is a finite subset of $V_n(R)$ such that $L(S) = V_n(R)$. In general if F is any field $V_n(F)$ is a finite dimensional vector space over F .

3. Let V be the set of all polynomials in $F[x]$ of degree $\leq n$. Let $S = \{1, x, x^2, \dots, x^n\}$. Then $L(S) = V$ and hence V is finite dimensional.

4. C is a finite dimensional vector space over R , since $L(\{1, i\}) = C$.

5. In $M_2(R)$ consider the set S consisting of the matrices

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; \quad C = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \quad D = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \text{ Then}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = aA + bB + cC + dD. \text{ Hence } L(S) = M_2(R) \text{ so that } M_2(R) \text{ is finite dimensional.}$$

Note:

All the vector spaces we have considered above are finite dimensional. However there are vector spaces which cannot be spanned by a finite number of vectors. For example, consider $R[x]$. Let S be any finite subset of $R[x]$. Let f be a polynomial of maximum degree in S . Let $\deg f = n$. Then any element of $L(S)$ is a polynomial of degree $\leq n$ and hence $L(S) \neq R[x]$. Thus $R[x]$ is not finite dimensional.

Throughout the rest of this chapter all the vector spaces we consider are finite dimensional.

Although we have defined what is meant by a finite dimensional space we have not yet defined what is meant by the dimension of a vector space. We now proceed to introduce the concepts necessary to define the dimension of a finite dimensional vector space.

Consider the vectors $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$ in $V_3(R)$.

Suppose that $\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = 0$.

Then $(\alpha_1, 0, 0) + (0, \alpha_2, 0) + (0, 0, \alpha_3) = (0, 0, 0)$.

$$\therefore (\alpha_1, \alpha_2, \alpha_3) = (0, 0, 0).$$

$$\therefore \alpha_1 = \alpha_2 = \alpha_3 = 0$$

(i.e) $\alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 = 0$ iff $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

Thus a linear combination of the vectors, e_1, e_2 and e_3 will yield the zero vector iff all the coefficients are zero.

Definition: 7.2.2

Let V be a vector space over a field F . A finite set of vectors $v_1, v_2, \dots, v_n$ in V is said to be **linearly independent** if

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$$

$$\Rightarrow \alpha_1 = \alpha_2 = \dots = \alpha_n = 0.$$

If $v_1, v_2, \dots, v_n$ are not linearly independent, then they are said to be **linearly dependent**.

Note:

If $v_1, v_2, \dots, v_n$ are linearly dependent, then there exist scalars $\alpha_1, \alpha_2, \dots, \alpha_n$ not all zero, such that $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$.

Examples:

1. In $V_n(F)$, $\{e_1, e_2, \dots, e_n\}$ is a linearly independent set of vectors, for, $\alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n = 0$.

$$\Rightarrow \alpha_1 (1, 0, \dots, 0) + \alpha_2 (0, 1, \dots, 0) + \dots + \alpha_n (0, 0, \dots, 1) = (0, 0, \dots, 0)$$

$$\Rightarrow (\alpha_1, \alpha_2, \dots, \alpha_n) = (0, 0, \dots, 0)$$

$$\Rightarrow \alpha_1 = \alpha_2 = \dots = \alpha_n = 0.$$

2. In $V_3(R)$ the vectors $(1, 2, 1), (2, 1, 0)$ and $(1, -1, 2)$ are linearly independent. For, let $\alpha_1 (1, 2, 1) + \alpha_2 (2, 1, 0) + \alpha_3 (1, -1, 2) = (0, 0, 0)$.

$$\therefore (\alpha_1 + 2\alpha_2 + \alpha_3, 2\alpha_1 + \alpha_2 - \alpha_3, \alpha_1 + 2\alpha_3) = (0, 0, 0).$$

$$\therefore \alpha_1 + 2\alpha_2 + \alpha_3 = 0 \dots \dots \dots (1)$$

$$2\alpha_1 + \alpha_2 - \alpha_3 = 0 \dots \dots \dots (2)$$

$$\alpha_1 + 2\alpha_3 = 0 \dots \dots \dots (3)$$

Solving equations (1), (2) and (3) we get $\alpha_1 = \alpha_2 = \alpha_3 = 0$.

$\therefore$ The given vectors are linearly independent.

3. In $V_3(\mathbb{R})$ the vectors $(1,4,-2)$, $(-2,1,3)$ and $(-4,11,5)$ are linearly dependent. For let $\alpha_1(1,4,-2) + \alpha_2(-2,1,3) + \alpha_3(-4,11,5) = (0,0,0)$

$$\therefore \alpha_1 - 2\alpha_2 - 4\alpha_3 = 0 \dots\dots\dots (1)$$

$$4\alpha_1 + \alpha_2 + 11\alpha_3 = 0 \dots\dots\dots (2)$$

$$-2\alpha_1 + 3\alpha_2 + 5\alpha_3 = 0 \dots\dots\dots (3)$$

From (1) and (2),

$$\frac{\alpha_1}{-18} = \frac{\alpha_2}{-27} = \frac{\alpha_3}{9} = k(\text{say})$$

$$\therefore \alpha_1 = -18k, \alpha_2 = -27k, \alpha_3 = 9k.$$

These values of α_1, α_2 and α_3 , for any k satisfy (3) also.

Taking $k = 1$ we get

$$\alpha_1 = -18, \alpha_2 = -27, \alpha_3 = 9 \text{ as a non-trivial solution.}$$

Hence the three vectors are linearly dependent.

4. Let V be a vector space over a field F . Then any subset S of V containing the zero vector is linearly dependent.

Proof:

$$\text{Let } S = \{0, v_1, \dots, v_n\}$$

$$\text{Clearly } \alpha 0 + 0v_1 + 0v_2 + \dots + 0v_n = 0 \text{ where } \alpha \text{ is any element of } F.$$

Hence for any $\alpha \neq 0$, we get a non-trivial linear combination of vectors in S giving the zero vector. Hence S is linearly dependent.

Exercises:

- Determine whether the following sets of vector are linearly independent or linearly dependent in $V_3(\mathbb{R})$.
 - $\{(1,0,0), (0,1,0), (1,1,0)\}$.
 - $\{(1,2,3), (2,3,1)\}$.
 - $\{(1,2,3), (4,1,5), (-4,6,2)\}$.
 - $\{(0,0,0), (2,5,3), (-1,0,6)\}$.
 - $\{(1,0,0), (1,1,0), (1,1,1), (0,1,0)\}$.
- Determine whether the following sets of vectors are linearly independent or not.

- (a) $\{(1,1,0,0), (0,0,1,1), (1,0,0,4), (0,0,0,2)\}$ in $V_4(\mathbb{R})$.
- (b) $\{(2i,1,0), (2,-i,1), (0,1,+i,-i)\}$ in $V_3(\mathbb{C})$.
- (c) $\{(\pi,0,0), (0,1,0), (0,0,\sqrt{5})\}$ in $V_3(\mathbb{R})$.
- (d) $V =$ the set of all polynomials of degree $\leq n$ in $\mathbb{R}[x]$ and

$$S = \{1, x, x^2, \dots, x^n\}.$$

(e) $\left\{ \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix} \right\}$ in $M_2(\mathbb{R})$.

3. In $V_3(\mathbb{Z}_5)$ determine whether the following sets of vectors are linearly dependent.

(a) $\{(1,3,2), (2,1,3)\}$

(b) $\{(1,1,2), (2,1,0), (0,4,1)\}$

4. In $V_2(\mathbb{R})$ prove that the vectors (a, b) and (c, d) are linearly dependent iff $ad - bc = 0$.

5. Let $\{v_1, v_2, v_3\}$ be a linearly independent set of vectors in $V_3(\mathbb{R})$. Show that

(a) $\{v_1 + v_2, v_2 + v_3, v_3 + v_1\}$ is linearly independent.

(b) $\{2v_1 + v_2, v_1 + v_2, v_1 - v_3\}$ is linearly independent.

6. If the vectors $(0, 1, a)$, $(1, a, 1)$ and $(a, 1, 0)$ of $V_3(\mathbb{R})$ are linearly dependent then find the value of a.

Answers:

- (b) is linearly independent.
- (a), (b), (c), (d), (e) are linearly independent.
- (a) is linearly independent 6. $a = 0, \pm\sqrt{2}$.

Theorem: 7.3

Any subset of a linearly independent set is linearly independent.

Proof:

Let V be a vector space over a field F .

Let $S = \{v_1, v_2, \dots, v_n\}$ be a linearly independent set.

Let S' be a subset of S . Without loss of generality we take

$$S' = \{v_1, v_2, \dots, v_k\} \text{ where } k \leq n.$$

Suppose S' is a linearly dependent set. Then there exist $\alpha_1, \alpha_2, \dots, \alpha_k$ in F not all zero, such that $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k = 0$.
Hence $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_k v_k + 0v_{k+1} + \dots + 0v_n = 0$ is a non-trivial linear combination giving the zero vector.

Here S is a linearly dependent set which is a contradiction.

Hence S' is linearly independent.

Theorem: 7.4

Any set containing a linearly dependent set is also linearly dependent.

Proof:

Let V be a vector space. Let S be a linearly dependent set. Let $S' \supset S$.

If S' is linearly independent S is also linearly independent (by theorem 7.3) which is a contradiction. Hence S' is a linearly dependent.

Theorem: 7.5

Let $S = \{v_1, v_2, \dots, v_n\}$ be a linearly independent set of vectors in a vector space V over a field F . Then every element of $L(S)$ can be uniquely written in the form

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n, \text{ where } \alpha_i \in F.$$

Proof:

By definition every elements of $L(S)$ is of the form

$$\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n.$$

Now, let $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = \beta_1 v_1 + \beta_2 v_2 + \dots + \alpha_n v_n$.

$$\text{Hence } (\alpha_1 - \beta_1)v_1 + (\alpha_2 - \beta_2)v_2 + \dots +$$

$$(\alpha_n - \beta_n)v_n = 0.$$

Since S is a linearly independent set, $\alpha_i - \beta_i = 0$ for all i .

$$\therefore \alpha_i = \beta_i \text{ for all } i. \text{ Hence the theorem.}$$

Theorem: 7.6

$S = \{v_1, v_2, \dots, v_n\}$ is a linearly dependent set of vectors in V iff there exists a vector $v_k \in S$ such that v_k is a linear combination of the preceding vectors $v_1, v_2, \dots, v_{k-1}$.

Proof:

Suppose $v_1, v_2, \dots, v_n$ are linearly dependent.

Then there exist $\alpha_1, \alpha_2, \dots, \alpha_n \in F$, not all zero, such that
 $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$.

Let k be the largest integer for which $\alpha_k \neq 0$.

Then $\alpha_1 v_1 + \dots + \alpha_k v_k = 0$.

$$\therefore \alpha_k v_k = -\alpha_1 v_1 - \alpha_2 v_2 - \dots - \alpha_{k-1} v_{k-1}.$$

$$\therefore v_k = (-\alpha_k^{-1} \alpha_1) v_1 + \dots + (-\alpha_k^{-1} \alpha_{k-1}) v_{k-1}.$$

$\therefore v_k$ is a linear combination of the preceding vectors.

Conversely, suppose there exists a vector v_k such that

$$v_k = \alpha_1 v_1 + \dots + \alpha_{k-1} v_{k-1}.$$

$$\text{Hence } -\alpha_1 v_1 - \dots - \alpha_{k-1} v_{k-1} + v_k + 0v_{k+1} + \dots + 0v_n = 0.$$

Since the coefficient of $v_k = 1$, we have

$S = \{v_1, \dots, v_n\}$ is linearly dependent.

Examples:

In $V_3(\mathbb{R})$, let $S = \{(1,0,0), (0,1,0), (0,0,1), (1,1,1)\}$

$$\text{Here } (1,1,1) = (1,0,0) + (0,1,0) + (0,0,1).$$

Thus $(1,1,1)$ is a linear combination of the preceding vectors.

Hence S is a linearly dependent set.

Theorem: 7.7

Let V be a vector space over F . Let $S = \{v_1, v_2, \dots, v_n\}$ and $L(S) = W$. Then there exists a linearly independent subset S' of S such that $L(S') = W$.

Proof:

Let $S = \{v_1, v_2, \dots, v_n\}$.

If S is linearly independent there is nothing to prove.

If not, let v_k be the first vector in S which is a linear combination of the preceding vectors.

Let $S_1 = \{v_1, v_2, \dots, v_{k-1}, v_{k+1}, \dots, v_n\}$.

(ie) S_1 is obtained by deleting the vector v_k from S . We claim that $L(S_1) = L(S) = W$.

Since $S_1 \subseteq S$, $L(S_1) \subseteq L(S)$. (refer theorem 7.2). Not, let $v \in L(S)$.

Then $v = \alpha_1 v_1 + \dots + \alpha_k v_k + \dots + \alpha_n v_n$.

Now, v_k is a linear combination of the preceding vectors.

Let $v_k = \beta_1 v_1 + \dots + \beta_{k-1} v_{k-1}$.

Hence

$v = \alpha_1 v_1 + \dots + \alpha_{k-1} v_{k-1} + \alpha_k (\beta_1 v_1 + \dots + \beta_{k-1} v_{k-1}) + \alpha_{k+1} v_{k+1} + \dots + \alpha_n v_n$
 $\therefore v$ can be expressed as a linear combination of the vectors of

S_1 so that $v \in L(S_1)$. Hence $L(S) \subseteq L(S_1)$

Thus $L(S) = L(S_1) = W$.

Now, if S_1 is linearly independent, the proof is complete.

If not, we continue the above process of removing a vector from S_1 , which is a linear combination of the preceding vectors until we arrive at a linearly independent subset S' of S such that $L(S') = W$.

7.3 Basis and Dimension:

Definition: 7.3.1

A linearly independent subset S of a vector space V which spans the whole space V is called a **basis** of the vector space.

Theorem: 7.8

Any finite – dimensional vector space V contains a finite number of linearly independent vectors which span V . (ie) A finite dimensional vector space has a basis consisting a finite number of vectors.

Proof:

Since V is finite dimensional there exists a finite subset S of V such that $L(S) = V$. By theorem 7.7 this set S contains a linearly independent subset $S' = \{v_1, v_2, \dots, v_n\}$ such that $L(S') = L(S) = V$.

Hence S' is a basis for V .

Theorem: 7.9

Let V be a vector space over a field F . Then $S = \{v_1, v_2, \dots, v_n\}$ is a basis for V iff every element of V can be uniquely expressed as a linear combination of element of S .

Proof:

Let S be a basis for V .

Then by definition S is linearly independent and $L(S) = V$.

Hence by theorem 7.5 every element of V can be uniquely expressed as a linear combination of elements of S .

Conversely, suppose every element of V can be uniquely expressed as a linear combination of elements of S .

Clearly $L(S) = V$.

Now, let $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$.

Also, $0v_1 + 0v_2 + \dots + 0v_n = 0$.

Thus we have expressed 0 as a linear combination of vectors of S in two ways.

$\therefore$ By hypothesis $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$.

Hence S is linearly independent. Hence S is a basis.

Examples:

1. $S = \{(1,0,0), (0,1,0), (0,0,1)\}$ is a basis for $V_3(\mathbb{R})$ for,

$$(a, b, c) = a(1,0,0) + b(0,1,0) + c(0,0,1).$$

$\therefore$ Any vector (a, b, c) of $V_3(\mathbb{R})$ has been expressed uniquely as a linear combination of the elements of S and hence S is a basis for $V_3(\mathbb{R})$.

2. $S = \{e_1, e_2, \dots, e_n\}$ is a basis for $V_n(\mathbb{F})$. This is known as the **standard basis** for $V_n(\mathbb{F})$.

3. $S = \{(1,0,0), (0,1,0), (1,1,1)\}$ is a basis for $V_3(\mathbb{R})$.

Proof:

We shall show that any element (a, b, c) of $V_3(\mathbb{R})$ can be uniquely expressed as a linear combination of the vectors of S .

$$\text{Let } (a, b, c) = \alpha(1,0,0) + \beta(0,1,0) + \gamma(1,1,1)$$

$$\text{Then } \alpha + \gamma = a, \beta + \gamma = b, \gamma = c.$$

$$\text{Hence } \alpha = a - c \text{ and } \beta = b - c.$$

$$\text{Thus } (a, b, c) = (a - c)(1,0,0) + (b - c)(0,1,0) + c(1,1,1).$$

$\therefore S$ is a basis for $V_3(\mathbb{R})$

4. $S = \{1\}$ is a basis for the vector space $\mathbb{R}$ over $\mathbb{R}$.

5. $S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ is a basis for $M_2(\mathbb{R})$, since

any matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ can be uniquely written as

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

6. $\{1, i\}$ is a basis for the vector space $\mathbb{C}$ over $\mathbb{R}$.

7. Let V be the set of all polynomials of degree $\leq n$ in $\mathbb{R}[x]$. Then $\{1, x, x^2, \dots, x^n\}$ is a basis for V .

8. $\{(1,0), (i,0), (0,1), (0,i)\}$ is a basis, for the vector space $\mathbb{C} \times \mathbb{C}$ over $\mathbb{R}$, for $(a + ib, c + id) = a(1,0) + b(i,0) + c(0,1) + d(0,i)$.

9. $S = \{(1,0,0), (0,1,0), (1,1,1), (1,1,0)\}$ spans the vectors space $V_3(\mathbb{R})$ but is not a basis.

Proof:

Let $S' = \{(1,0,0), (0,1,0), (1,1,1)\}$. Then $L(S') = V_3(\mathbb{R})$ (refer example 3).

Now, since $S \subseteq S'$, we get $L(S) = V_3(\mathbb{R})$. Thus S spans $V_3(\mathbb{R})$.

But S is linearly dependent since $(1,1,0) = (1,0,0) + (0,1,0)$.

Hence S is not a basis.

10. $S = \{(1,0,0), (1,1,0)\}$ is linearly independent but not a basis of $V_3(\mathbb{R})$.

Proof:

Let $\alpha(1,0,0) + \beta(1,1,0) = (0,0,0)$. Then $\alpha + \beta = 0$ and $\beta = 0$.

$\therefore \alpha = \beta = 0$. Hence S is linearly independent.

Also $L(S) = \{(a,b,0) / a, b \in \mathbb{R}\} \neq V_3(\mathbb{R})$.

$\therefore S$ is not a basis.

Exercises:

1. Show that the following three vectors form a basis for $V_3(\mathbb{R})$.

(a) $(1,2,-3), (2,5,1), (-1,1,4)$

(b) $(1,1,0), (0,1,1), (1,0,1)$

- (c) $(2, -3, 1), (0, 1, 2), (1, 1, 2)$.
2. Show that the following sets of vectors do not form a basis for $V_3(\mathbb{R})$.
- (a) $\{(1, 0, 0), (1, 1, 0)\}$
- (b) $\{(1, 2, 1), (1, 3, 5), (-1, 0, 1), (1, -1, 2)\}$
- (c) $\{(0, 0, 0), (1, 0, 0), (0, 1, 0), (0, 0, 1)\}$
- (d) $\{(3, 2, 1), (3, 1, 5), (3, 4, -7)\}$
- (e) $\{(1, 2, 3), (2, 3, 4), (3, 4, 5)\}$
3. Show that $(1, i, 0), (2i, 1, 1), (0, 1 + i, 1 - i)$ form a basis for $V_3(\mathbb{C})$.
4. Find the basis for the vector space consisting of all matrices of the form
- (a) $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$
- (b) $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$
5. If $\{v_1, v_2, v_3\}$ is a basis for $V_3(\mathbb{R})$, Show that $\{v_1 + v_2, v_2 + v_3, v_3 + v_1\}$ is also a basis. Is this true in (a) $V_3(\mathbb{Z}_2)$ (b) $V_3(\mathbb{Z}_3)$?

Answers:

4. (a) $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$
- (b) $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$.

Theorem: 7.10

Let V be a vector space over a field F . Let $S = \{v_1, v_2, \dots, v_n\}$ span V . Let $S = \{w_1, w_2, \dots, w_n\}$ be a linearly independent set of vectors in V . Then $m \leq n$.

Proof:

Since $L(S) = V$, every vector in V and in particular w_1 , is a linear combination of $v_1, v_2, \dots, v_n$.

Hence $S_1 = \{w_1, v_1, v_2, \dots, v_n\}$ is a linearly dependent set of vectors. Hence there exists a vector $v_k \neq w_1$ in S_1 which is a linear combination of the preceding vectors.

Let $S_2 = \{w_1, v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n\}$.

Clearly, $L(S_2) = V$.

Hence w_2 is a linear combination of the vectors in S_2 .

Hence $S_3 = \{w_2, w_1, v_1, \dots, v_{k-1}, v_{k+1}, \dots, v_n\}$ is linearly dependent. Hence there exists a vector in S_3 which is a linear combination of the preceding vectors. Since the w_i 's are linearly independent, this vector cannot be w_2 or w_1 and hence must be some v_j where $j \neq k$ (say, with $j > k$).

Deletion of v_j from the set S_3 gives the set $S_4 = \{w_2, w_1, v_1, v_2, \dots, v_{k-1}, v_{k+1}, \dots, v_{j-1}, v_{j+1}, \dots, v_n\}$ of n vectors spanning V .

In this process, at each step we insert one vector from $\{w_1, w_2, \dots, w_m\}$ and delete one vector from $\{v_1, v_2, \dots, v_n\}$.

If $m > n$ after repeating this process n times, we arrive at the set $\{w_n, w_{n-1}, \dots, w_1\}$ which spans V .

Hence w_{n+1} is a linear combination of $w_1, w_2, \dots, w_n$. Hence $\{w_1, w_2, \dots, w_n, w_{n+1}, \dots, w_m\}$ is linearly dependent which is a contradiction.

Hence $m \leq n$.

Theorem: 7.11

Any two bases of a finite dimensional vector space V have the same number of elements.

Proof:

Since V is finite dimensional, it has a basis say $S = \{v_1, v_2, \dots, v_n\}$.

Let $S' = \{w_1, w_2, \dots, w_m\}$ be any other basis for V .

Now, $L(S) = V$ and S' is a set of m linearly independent vectors.

Hence by Theorem 7.10, $m \leq n$.

Also, since $L(S') = V$ and S is a set of n linearly independent vectors,

$n \leq m$. Hence $m = n$.

Definition: 7.3.2

Let V be a finite dimensional vector space over a field F . The number of elements in any basis of V is called the **dimension** of V and is denoted by $\dim V$.

Examples:

1. $\dim V_n(\mathbb{R}) = n$, since $\{e_1, e_2, \dots, e_n\}$ is a basis of $V_n(\mathbb{R})$.
2. $M_2(\mathbb{R})$ is a vector space of dimension 4 over $\mathbb{R}$ since $\left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$ is a basis for $M_2(\mathbb{R})$.
3. $\mathbb{C}$ is a vector space of dimension 2 over $\mathbb{R}$ since $\{1, i\}$ is a basis for $\mathbb{C}$.
4. Let V be the set of all polynomials of degree $\leq n$ in $\mathbb{R}[x]$. V is a vector space over $\mathbb{R}$ having dimension $n+1$, since $\{1, x, x^2, \dots, x^n\}$ is a basis for V .

Theorem: 7.12

Let V be a vector space of dimension n . Then

- (i) Any set of m vectors where $m > n$ is linearly dependent.
- (ii) Any set of m vectors where $m < n$ cannot span V .

Proof:

- (i) Let $S = \{v_1, v_2, \dots, v_n\}$ be a basis for V . Hence $L(S) = V$.

Let S' be any set consisting of m vectors where $m > n$.

Suppose S' is linearly independent. Since S spans V by Theorem 7.10, $m \leq n$ which is a contradiction.

Hence S' is linearly dependent.

- (ii) Let S' be a set consisting of m vectors where $m < n$. Suppose $L(S') = V$.

Now, $S = \{v_1, v_2, \dots, v_n\}$ is a basis for V and hence linearly independent. Hence by Theorem 7.10 $n \leq m$ which is a contradiction. Hence S' cannot span V .

Theorem: 7.13

Let V be a finite dimensional vector space over a field F . Any linearly independent set of vectors in V is part of a basis.

Proof:

Let $S = \{v_1, v_2, \dots, v_r\}$ be a linearly independent set of vectors.

If $L(S) = V$ then S itself is a basis.

If $L(S) \neq V$, choose an element $v_{r+1} \in V - L(S)$.

Now, consider $S_1 = \{v_1, v_2, \dots, v_r, v_{r+1}\}$.

We shall prove that S_1 is linearly independent by showing that no vector in S_1 is a linear combination of the preceding vectors. (refer theorem 7.6).

Since $\{v_1, v_2, \dots, v_r\}$ is linearly independent, v_i where $1 \leq i \leq r$ is not a linear combination of the preceding vectors.

Also $v_{r+1} \notin L(S)$ and hence v_{r+1} is not a linear combination of $v_1, v_2, \dots, v_r$.

Hence S_1 is linearly independent.

If $L(S_1) = V$, then S_1 is a basis for V . If not we take an element $v_{r+2} \in V - L(S_1)$ and proceed as before. Since the dimension of V is finite, this process must stop at a certain stage giving the required basis containing S .

Theorem: 7.14

Let V be a finite dimensional vector space over a field F . Let A be a subspace of V . Then there exists a subspace B of V such that $V = A \oplus B$.

Proof:

Let $S = \{v_1, v_2, \dots, v_r\}$ be a basis of A .

By theorem 7.13, we can find $w_1, w_2, \dots, w_s \in V$ such that $S' = \{v_1, v_2, \dots, v_r, w_1, w_2, \dots, w_s\}$ is a basis of V .

Now, let $B = L(\{w_1, w_2, \dots, w_s\})$

We claim that $A \cap B = \{0\}$ and $V = A + B$.

Now, let $v \in A \cap B$. Then $v \in A$ and $v \in B$.

$$\text{Hence } v = \alpha_1 v_1 + \dots + \alpha_r v_r,$$

$$= \beta_1 w_1 + \dots + \beta_s w_s,$$

$$\therefore \alpha_1 v_1 + \dots + \alpha_r v_r - \beta_1 w_1 - \dots - \beta_s w_s = 0.$$

Now, since S' is linearly independent, $\alpha_i = 0 = \beta_j$ for all i and j .

Hence $v = 0$. Thus $A \cap B = \{0\}$.

Now, let $v \in V$.

Then $v = \{\alpha_1 v_1 + \dots + \alpha_r v_r\} + \{\beta_1 w_1 + \dots + \beta_s w_s\} \in A + B$.

Hence $A + B = V$ so that $V = A \oplus B$.

Exercises:

1. Let V be a finite - dimensional vector space. Let A and B be subspaces of V such that $V = A \oplus B$. Then show that $\dim V = \dim A + \dim B$.

2. Construct 3 subspaces W_1, W_2, W_3 of a vector space V such that $V = W_1 \oplus W_2 = W_1 \oplus W_3$ but $W_2 \neq W_3$.

3. For each of the following subspaces A of $V_3(\mathbb{R})$ find another subspace B such that $A \oplus B = V_3(\mathbb{R})$

$$(i) \quad A = L\{(1,1,0), (0,1,1)\}$$

$$(ii) \quad A = L\{(1,1,1)\}$$

$$(iii) \quad A = L(\{e_1, e_2, e_3\}).$$

Definition: 7.3.3

Let V be a vector space and $S = \{v_1, v_2, \dots, v_n\}$ be a set of independent vectors in V . Then S is called a **maximal linearly independent set** if for every $v \in V - S$, the set $\{v, v_1, v_2, \dots, v_n\}$ is linearly dependent.

Definition:

Let $S = \{v_1, v_2, \dots, v_n\}$ be a set of vectors in V and let $L(S) = V$

Then S is called **minimal generating set** if for any $v_i \in S$.

$$L(S - \{v_i\}) \neq V.$$

Theorem: 7.15

Let V be a vector space over a field F . Let

Space for Hints

$S = \{v_1, v_2, \dots, v_n\} \subseteq V$. Then the following are equivalent.

- (i) S is a basis for V .
- (ii) S is a maximal linearly independent set.
- (iii) S is a minimal generating set.

Proof:

(i) $\Rightarrow$ (ii) Let $S = \{v_1, v_2, \dots, v_n\}$ be a basis for V . Then by theorem 7.12 any $n + 1$ vectors in V are linearly dependent and hence S is a maximal linearly independent set.

(ii) $\Rightarrow$ (i) Let $S = \{v_1, v_2, \dots, v_n\}$ be a maximal linearly independent set. Now to prove that S is a basis for V we shall show that $L(S) = V$.

Obviously $L(S) \subseteq V$.

Now, let $v \in V$.

If $v \in S$, then $v \in L(S)$. (since $S \subseteq L(S)$)

If $v \notin S$, $S' = \{v_1, v_2, \dots, v_n, v\}$ is a linearly dependent set (since S is a maximal linearly independent set)

$\therefore$ There exists a vector in S' which is a linear combination of the preceding vectors.

Since $v_1, v_2, \dots, v_n$ are linearly independent, this vector must be v . Thus v is a linear combination of $v_1, v_2, \dots, v_n$. Therefore $v \in L(S)$.

Hence $V \subseteq L(S)$. Thus $V = L(S)$.

(i) $\Rightarrow$ (iii) Let $S = \{v_1, v_2, \dots, v_n\}$ be a basis. Then $L(S) = V$.

If S is not minimal, there exists $v_i \in S$ such that $L(S - \{v_i\}) = V$.

Since S is linearly independent, $S - \{v_i\}$ is also linearly independent. Thus $S - \{v_i\}$ is a basis consisting of $n - 1$ elements which is a contradiction.

Hence S is a minimal generating set.

(iii) $\Rightarrow$ (i) Let $S = \{v_1, v_2, \dots, v_n\}$ be a minimal generating set. To prove that S is a basis, we have to show that S is linearly independent.

If S is linearly dependent, there exists a vector v_k which is a linear combination of the preceding vectors.

Clearly $L(S - \{v_k\}) = V$ contradicting the minimality of S .

Thus S is linearly independent and since $L(S) = V$, S is a basis for V .

Theorem: 7.16

Any vector space of dimension n over a field F is isomorphic to $V_n(F)$.

Proof:

Let V be a vector space of dimension n . Let $\{v_1, v_2, \dots, v_n\}$ be a basis for V .

Then we know that if $v \in V$, v can be written uniquely as $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$, where $\alpha_i \in F$.

Now, consider the map $f: V \rightarrow V_n(F)$ given by

$$f(\alpha_1 v_1 + \dots + \alpha_n v_n) = (\alpha_1, \alpha_2, \dots, \alpha_n).$$

Clearly f is 1-1 and onto.

Let $v, w \in V$.

Then $v = \alpha_1 v_1 + \dots + \alpha_n v_n$ and

$$w = \beta_1 v_1 + \dots + \beta_n v_n.$$

$$\begin{aligned} f(v+w) &= f[(\alpha_1 + \beta_1)v_1 + \dots + (\alpha_n + \beta_n)v_n] = ((\alpha_1 + \beta_1), (\alpha_2 + \beta_2), \dots, (\alpha_n + \beta_n)) \\ &= (\alpha_1, \alpha_2, \dots, \alpha_n) + (\beta_1, \beta_2, \dots, \beta_n) \\ &= f(v) + f(w). \end{aligned}$$

Also $f(\alpha v) = f(\alpha \alpha_1 v_1 + \dots + \alpha \alpha_n v_n)$

$$= (\alpha \alpha_1, \alpha \alpha_2, \dots, \alpha \alpha_n)$$

$$= \alpha(\alpha_1, \alpha_2, \dots, \alpha_n)$$

$$= \alpha f(v).$$

Hence f is an isomorphism of V to $V_n(F)$.

Corollary:

Any two vector spaces of the same dimension over a field F are isomorphic, for, if the vector spaces are of dimension n , each is isomorphic to $V_n(F)$ and hence they are isomorphic.

Theorem: 7.17

Let V and W be vector spaces over a field F . Let $T: V \rightarrow W$ be an isomorphism. Then T maps a basis of V onto a basis of W .

Proof:

Let $\{v_1, v_2, \dots, v_n\}$ be a basis for V .

We shall prove that $T(v_1), T(v_2), \dots, T(v_n)$ are linearly independent and that they span W .

$$\text{Now, } \alpha_1 T(v_1) + \alpha_2 T(v_2) + \dots + \alpha_n T(v_n) = 0.$$

$$\Rightarrow T(\alpha_1 v_1) + T(\alpha_2 v_2) + \dots + T(\alpha_n v_n) = 0.$$

$$\Rightarrow T(\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n) = 0.$$

$$\Rightarrow \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0 \quad (\text{since } T \text{ is } 1-1)$$

$$\Rightarrow \alpha_1 = \alpha_2 = \dots = \alpha_n = 0$$

(since $v_1, v_2, \dots, v_n$ are linearly independent).

$\therefore T(v_1), T(v_2), \dots, T(v_n)$ are linearly independent.

Now, let $w \in W$. Then since T is onto, there exists a vector $v \in V$ such that $T(v) = w$.

$$\text{Let } v = \alpha_1 v_1 + \dots + \alpha_n v_n.$$

$$\text{Then } w = T(v)$$

$$= T(\alpha_1 v_1 + \dots + \alpha_n v_n).$$

$$= \alpha_1 T(v_1) + \dots + \alpha_n T(v_n).$$

Thus w is a linear combination of the vectors.

$$T(v_1), \dots, T(v_n).$$

$\therefore T(v_1), \dots, T(v_n)$ span W and hence is a basis for W .

Corollary:

Two finite dimensional vector spaces V and W over a field F are isomorphic iff they have the same dimension.

Theorem: 7.18

Let V and W be finite dimensional vector spaces over a field F . Let $\{v_1, v_2, \dots, v_n\}$ be a basis for V and let $w_1, w_2, \dots, w_n$ be any n vectors in W (not necessarily distinct). Then there exists a unique linear transformation $T: V \rightarrow W$ such that $T(v_i) = w_i, i = 1, 2, \dots, n$.

Proof:

$$\text{Let } v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n \in V.$$

We define $T(v) = \alpha_1 w_1 + \alpha_2 w_2 + \dots + \alpha_n w_n$.

Now, let $x, y \in V$.

Let $x = \alpha_1 v_1 + \dots + \alpha_n v_n$ and

$y = \beta_1 v_1 + \dots + \beta_n v_n$.

$$\therefore x + y = (\alpha_1 + \beta_1)v_1 + \dots + (\alpha_n + \beta_n)v_n.$$

$$\begin{aligned} \therefore T(x + y) &= (\alpha_1 + \beta_1)w_1 + \dots + (\alpha_n + \beta_n)w_n \\ &= (\alpha_1 w_1 + \dots + \alpha_n w_n) + (\beta_1 w_1 + \dots + \beta_n w_n) \\ &= T(x) + T(y). \end{aligned}$$

Similarly $T(\alpha x) = \alpha T(x)$.

Hence T is a linear transformation.

Also $v_1 = 1v_1 + 0v_2 + \dots + 0v_n$.

Hence $T(v_1) = 1w_1 + 0w_2 + 0w_n = w_1$.

Similarly $T(v_i) = w_i$ for all $i = 1, 2, \dots, n$.

Now, to prove the uniqueness, let $T': V \rightarrow W$ be any other linear transformation such that $T'(v_i) = w_i$.

Let $v = \alpha_1 v_1 + \dots + \alpha_n v_n \in V$

$$\begin{aligned} T'(v) &= \alpha_1 T'(v_1) + \dots + \alpha_n T'(v_n) \\ &= \alpha_1 w_1 + \dots + \alpha_n w_n = T(v). \end{aligned}$$

Hence $T = T'$.

Remark:

The above theorem shows that a linear transformation is completely determined by its values on the elements of a basis.

Theorem: 7.19

Let V be a finite dimensional vector space over a field F . Let W be a subspace of V . Then

$$(i) \quad \dim W \leq \dim V.$$

$$(ii) \quad \dim \frac{V}{W} = \dim V - \dim W.$$

Proof:

- (i) Let $S = \{w_1, w_2, \dots, w_m\}$ be a basis for W . Since W is a subspace of V , S is a part of a basis for V .

Space for Hints

Hence $\dim W \leq \dim V$.

(ii) Let $\dim V = n$ and $\dim W = m$.

Let $S = \{w_1, w_2, \dots, w_m\}$ be a basis for W . Clearly S is a linearly independent set of vectors in V .

Hence S is a part of a basis in V . Let $\{w_1, w_2, \dots, w_m, v_1, v_2, \dots, v_r\}$ be a basis for V . Then $m + r = n$.

Now, we claim $S' = \{W + v_1, W + v_2, \dots, W + v_r\}$ is a basis for $\frac{V}{W}$.

$$\begin{aligned}\alpha_1(W + v_1) + \alpha_2(W + v_2) + \dots + \alpha_r(W + v_r) &= W + 0 \\ \Rightarrow (W + \alpha_1 v_1) + (W + \alpha_2 v_2) + \dots + (W + \alpha_r v_r) &= W \\ \Rightarrow \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_r v_r &\in W\end{aligned}$$

Now, since $\{w_1, w_2, \dots, w_m\}$ is a basis for W

$$\alpha_1 v_1 + \dots + \alpha_r v_r = \beta_1 w_1 + \dots + \beta_m w_m.$$

$$\therefore \alpha_1 v_1 + \dots + \alpha_r v_r - \beta_1 w_1 - \dots - \beta_m w_m = 0.$$

$$\therefore \alpha_1 = \alpha_2 = \dots = \alpha_r = \beta_1 = \beta_2 = \dots = \beta_m = 0$$

$\therefore S'$ is a linearly independent set.

Now, let $W + v \in \frac{V}{W}$.

$$\text{Let } v = \alpha_1 v_1 + \dots + \alpha_r v_r + \beta_1 w_1 + \dots + \beta_m w_m.$$

$$\begin{aligned}\text{Then } W + v &= W + (\alpha_1 v_1 + \dots + \alpha_r v_r + \beta_1 w_1 + \dots + \beta_m w_m) \\ &= W + (\alpha_1 v_1 + \dots + \alpha_r v_r) \\ &\quad (\text{since } \beta_1 w_1 + \dots + \beta_m w_m \in W) \\ &= (W + \alpha_1 v_1) + \dots + (W + \alpha_r v_r) \\ &= \alpha_1(W + v_1) + \dots + \alpha_r(W + v_r).\end{aligned}$$

Hence S' spans $\frac{V}{W}$ so that S' is a basis for $\frac{V}{W}$.

$$\begin{aligned}\therefore \dim \frac{V}{W} &= r = n - m \\ &= \dim V - \dim W.\end{aligned}$$

Theorem: 7.20

Let V be a finite – dimensional vector space over a field F . Let A and B be subspace of V .

Then $\dim(A + B) = \dim A + \dim B - \dim(A \cap B)$

Proof:

A and B are subspaces of V. Hence $A \cap B$ is subspace of V.

Let $\dim(A \cap B) = r$.

Let $S = \{v_1, v_2, \dots, v_r\}$ be a basis for $A \cap B$.

Since $A \cap B$ is a subspace of A and B, S is a part of a basis for A and B.

Let $\{v_1, v_2, \dots, v_r, u_1, u_2, \dots, u_s\}$ be a basis for A and

$\{v_1, v_2, \dots, v_r, w_1, w_2, \dots, w_t\}$ be a basis for B.

We shall prove that $S' = \{v_1, \dots, v_r, u_1, \dots, u_s, w_1, \dots, w_t\}$ is a basis for $A + B$.

Let

$$\alpha_1 v_1 + \dots + \alpha_r v_r + \beta_1 u_1 + \dots + \beta_s u_s + \gamma_1 w_1 + \dots + \gamma_t w_t = 0.$$

Then

$$\beta_1 u_1 + \dots + \beta_s u_s = -(\gamma_1 w_1 + \dots + \gamma_t w_t) - (\alpha_1 v_1 + \dots + \alpha_r v_r) \in B$$

Hence $\beta_1 u_1 + \dots + \beta_s u_s \in B$.

Also $\beta_1 u_1 + \dots + \beta_s u_s \in A$.

Hence $\beta_1 u_1 + \dots + \beta_s u_s \in A \cap B$.

$$\therefore \beta_1 u_1 + \dots + \beta_s u_s = \delta_1 v_1 + \dots + \delta_r v_r.$$

$$\therefore \beta_1 u_1 + \dots + \beta_s u_s - \delta_1 v_1 - \dots - \delta_r v_r = 0.$$

$$\therefore \beta_1 = \dots = \beta_s = \delta_1 = \dots = \delta_r = 0$$

(since $\{u_1, \dots, u_s, v_1, \dots, v_r\}$ is linearly independent)

Similarly we can prove $\gamma_1 = \gamma_2 = \dots = \gamma_t = 0$.

$$\therefore \alpha_i = \beta_j = \gamma_k = 0 \text{ for } 1 \leq i \leq r, 1 \leq j \leq s; 1 \leq k \leq t$$

Thus S' is a linearly independent set.

Clearly S' spans $A + B$.

$\therefore S'$ is a basis for $A + B$.

Hence $\dim(A + B) = r + s + t$.

Also $\dim A = r + s; \dim B = r + t$ and $\dim(A \cap B) = r$.

$$\begin{aligned}\therefore \dim A + \dim B - \dim A \cap B &= (r+s) + (r+t) - r \\ &= r+s+t \\ &= \dim(A+B).\end{aligned}$$

Aliter:

$$\text{By theorem 6.7, } \frac{A+B}{A} = \frac{B}{A \cap B}.$$

$$\text{Hence } \dim\left(\frac{A+B}{A}\right) = \dim\left(\frac{B}{A \cap B}\right).$$

$$\therefore \dim(A+B) - \dim A = \dim B - \dim(A \cap B).$$

$$\therefore \dim(A+B) = \dim A + \dim B - \dim(A \cap B).$$

Corollary:

$$\text{If } V = A \oplus B, \dim V = \dim A + \dim B.$$

Proof:

$$V = A \oplus B \Rightarrow A+B = V \text{ and } A \cap B = \{0\}$$

$$\therefore \dim(A \cap B) = 0.$$

$$\text{Hence } \dim V = \dim A + \dim B.$$

Exercises:

1. Find the dimension of the subspace spanned by the following vectors in $V_3(\mathbb{R})$.

(a) $(1,1,1), (-1,-1,-1).$

(b) $(1,0,2), (2,0,1), (1,0,1)$

(c) $(1,2,-3), (0,0,1), (-1,2,1).$

(d) $(1,1,2), (-1,1,0).$

2. Find the dimension of the subspace spanned by the following vectors in $V_4(\mathbb{R})$.

(a) e_1, e_2, e_3, e_4

(b) e_1, e_2

(c) e_1, e_2, e_3

(d) e_1

3. In $V_3(\mathbb{R})$, find $\dim(A+B)$ and $\dim(A \cap B)$ where

- (a) A is the subspace spanned by $(1,1,1)$ and B is the subspace spanned by $(-1,-1,-1)$
- (b) A is the subspace spanned by $(1,1,1)$ and B is the subspace spanned by $(1,2,1)$.
- (c) A is the subspace spanned by $(1,1,1)$ and $(1,2,1)$ and B is the subspace spanned by $(0,0,1)$.
- (d) A is the subspace spanned by $(1,1,1)$ and $(1,2,1)$ and B is the subspace spanned by $(1,-1,1)$ and $(-1,1,-1)$.

4. Let V_1 and V_2 be subspaces of V such that $V_1 \cap V_2$ is the zero space.

Prove that $\dim V_1 + \dim V_2 \leq \dim V$.

5. Let V_1 and V_2 be subspaces of V such that every vector $v \in V$ can be represented as $v = v_1 + v_2$ where $v_1 \in V_1$ and $v_2 \in V_2$.

Prove that $\dim V_1 + \dim V_2 \geq \dim V$.

6. If A and B are finite dimensional subspaces of V such that $A \subseteq B$ and $\dim A = \dim B$ then show that $A = B$.

7. Let S be a subspace of finite – dimensional vector space V . If $\dim V = \dim S$ then prove that $S = V$.

8. Let W_1 and W_2 be two subspaces of a finite dimensional vector space V .

If $\dim V = \dim W_1 + \dim W_2$ and $W_1 \cap W_2 = \{0\}$ prove that,

$$V = W_1 \oplus W_2.$$

Answers:

- | | | | |
|-------------|----------|----------|----------|
| 1. (a) 1 | (b) 2 | (c) 3 | 1) 2 |
| 2. (a) 4 | (b) 2 | (c) 3 | (d) 1 |
| 3. (a) 1; 1 | (b) 2; 0 | (c) 3; 0 | (d) 3; 0 |

7.4 Rank and Nullity

Definition: 7.4.1

Let $T: V \rightarrow W$ be a linear transformation. Then the dimension of $T(V)$ is called the **rank** of T . The dimension of $\ker T$ is called the **nullity** of T .

Space for Hints

Theorem: 7.21

Let $T: V \rightarrow W$ be a linear transformation. Then $\dim V = \text{rank } T + \text{nullity } T$.

Proof:

We know that $V / \ker T = T(V)$.

$$\therefore \dim V - \dim(\ker T) = \dim(T(V))$$

$$\therefore \dim V - \text{nullity } T = \text{rank } T$$

$$\therefore \dim V = \text{nullity } T + \text{rank } T$$

Note:

$\ker T$ is also called null space of T .

Example:

Let V denote the set of all polynomials of degree $\leq n$ in $R[x]$. Let

$T: V \rightarrow V$ be defined by $T(f) = \frac{df}{dx}$. We know that T is a linear

transformation. Since $\frac{df}{dx} = 0 \Leftrightarrow f$ is constant, $\ker T$ consists of all constant

polynomials. The dimension of this subspace of V is 1. Hence $\text{nullity } T$ is 1.

Since $\dim V = n + 1$, $\text{rank } T = n$.

Exercises:

1. Find the rank and nullity of the linear transformations given in section 6.3.
2. Let V be a finite - dimensional vector space over a field F . Let $T: V \rightarrow V$ be a linear transformation such that $\text{rank } T = \text{nullity } T$. Show that $\dim V$ is even. Give an example of such a transformation.

Answers:

1. 1. $\text{nullity } T = \dim V$; $\text{rank } T = 0$.
2. $\text{nullity } T = 0$; $\text{rank } T = \dim V$;
3. $\text{nullity } T = \dim W$;
 $\text{rank } T = \dim V - \dim W$.
4. $\text{nullity } T = 2$; $\text{rank } T = 1$;
5. $\text{nullity } T = 1$; $\text{rank } T = n$.
6. $\text{nullity } T = 0$; $\text{rank } T = n + 1$.

Definition: 7.4.2**Space for Hints**

A linear transformation $T : V \rightarrow W$ is called **non – singular** if T is 1 – 1; otherwise T is called **singular**.

Exercises:

1. Let V and W be finite dimensional vector spaces over a field F and $\dim V > \dim W$. Then show that any linear transformation $T : V \rightarrow W$ is singular.
2. Let V be a finite – dimensional vector space over a field F . Then any non – singular linear transformation $T : V \rightarrow V$ is onto.
3. Let $T : V \rightarrow W$ be a linear transformation. Show that T is a non – singular iff $\text{rank } T = \dim V$.
4. Let $T_1 : V \rightarrow V$ and $T_2 : V \rightarrow V$ be linear transformations. Prove that
 - (a) $\text{rank } (T_2 T_1) \leq \text{rank } T_2$.
 - (b) $\text{nullity } (T_2 T_1) \geq \text{nullity } T_1$.
 - (c) $\text{rank } (T_2 T_1) = \text{rank } T_2$ iff T_1 is non – singular.
5. Let $T : V \rightarrow W$ be a linear transformation which is both 1 – 1 and onto. Show that $T^{-1} : W \rightarrow V$ is a linear transformation.
6. Determine which of the following statements are true and which are false.
 - (a) If $T : V \rightarrow W$ is a linear transformation then
 - (i) $\text{rank } T \leq \dim V$.
 - (ii) $\text{nullity } T \leq \dim V$.
 - (iii) $\text{Rank } T \leq \dim W$.
 - (iv) If T is onto $\text{rank } T = \dim W$.
 - (v) If T is non – singular $\text{rank } T = \dim V$
 - (vi) $\text{rank } T = \dim V \Rightarrow \text{nullity } T = 0$.
 - (b) Every linear transformation $T : V_4(\mathbb{R}) \rightarrow V_3(\mathbb{R})$ is singular.
 - (c) If $T : V \rightarrow W$ is non – singular and $\{v_1, \dots, v_n\}$ is a basis then $\{T(v_1), \dots, T(v_n)\}$ is a basis for W .

Answers:

6. (a) (i) T (ii) T (iii) T (iv) T (v) T (vi) T

7.5.0 Introduction

Let $T: V \rightarrow W$ be a linear transformation. We have seen that T is completely specified by the elements $T(v_1), T(v_2), \dots, T(v_m)$. Now, let

$$\begin{aligned} T(\mathbf{v}_1) &= a_{11}\mathbf{w}_1 + a_{12}\mathbf{w}_2 + \dots + a_{1n}\mathbf{w}_n \\ T(\mathbf{v}_2) &= a_{21}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \dots + a_{2n}\mathbf{w}_n \\ &\vdots \\ T(\mathbf{v}_m) &= a_{m1}\mathbf{w}_1 + a_{m2}\mathbf{w}_2 + \dots + a_{mn}\mathbf{w}_n \end{aligned} \quad (1)$$

Hence $T(v_1), T(v_2), \dots, T(v_m)$ are completely specified by the mn elements a_{ij} of the field F . These a_{ij} can be conveniently arranged in the form of m rows and n columns as follows.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Such an array of mn elements of F arranged in m rows and n columns is known as $m \times n$ matrix over the field F and is denoted by (a_{ij}) . Thus to every linear transformation T there is associated with it an $m \times n$ matrix over F . Conversely any $m \times n$ matrix over F . Conversely any $m \times n$ matrix over F defines a linear transformation $T : V \rightarrow W$ given by the formula (1).

Note:

The $m \times n$ matrix which we have associated with a linear transformation $T: V \rightarrow W$ depends on the choice of the basis for V and W .

For example, consider the linear transformation $T: V_2(\mathbb{R}) \rightarrow V_2(\mathbb{R})$ given by $T(a, b) = (a, a + b)$. Choose $\{e_1, e_2\}$ as a basis both for the domain and the range.

Then $T(e_1) = (1,1) = e_1 + e_2$

$$T(e_2) = (0,1) = e_2.$$

Hence the matrix representing T is $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$

Now, we choose $\{e_1, e_2\}$ as a basis for the domain and $\{(1,1), (1,-1)\}$ as a basis for the range.

Let $w_1 = (1,1)$ and $w_2 = (1,-1)$.

Then $T(e_1) = (1,1) = w_1$, and $T(e_2) = (0,1) = (1/2)w_1 - (1/2)w_2$.

Hence the matrix representing T is $\begin{pmatrix} 1 & 0 \\ 1/2 & -1/2 \end{pmatrix}$

Problems:

Problem: 1

Obtain the matrix representing the linear transformation $T: V_3(\mathbb{R}) \rightarrow V_3(\mathbb{R})$ given by $T(a,b,c) = (3a, a-b, 2a+b+c)$ w.r.t. the standard basis $\{e_1, e_2, e_3\}$.

Solution:

$$T(e_1) = T(1,0,0) = (3,1,2) = 3e_1 + e_2 + 2e_3$$

$$T(e_2) = T(0,1,0) = (0,-1,1) = -e_2 + e_3$$

$$T(e_3) = T(0,0,1) = (0,0,1) = e_3$$

Thus the matrix representing T is $\begin{pmatrix} 3 & 1 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$

Problem: 2

Find the linear transformation $T: V_3(\mathbb{R}) \rightarrow V_3(\mathbb{R})$ determined by

the matrix $\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ -1 & 3 & 4 \end{pmatrix}$ w.r.t the standard basis $\{e_1, e_2, e_3\}$.

Solution:

$$T(e_1) = e_1 + 2e_2 + e_3 = (1,2,1).$$

$$T(e_2) = 0e_1 + e_2 + e_3 = (0,1,1)$$

$$T(e_3) = -e_1 + 3e_2 + 4e_3 = (-1,3,4)$$

Now, $(a, b, c) = a(1, 0, 0) + b(0, 1, 0) + c(0, 0, 1)$

$$= ae_1 + be_2 + ce_3.$$

$$\therefore T(a, b, c) = T(ae_1 + be_2 + ce_3)$$

$$= aT(e_1) + bT(e_2) + cT(e_3)$$

$$= a(1, 2, 1) + b(0, 1, 1) + c(-1, 3, 4).$$

$$\therefore T(a, b, c) = (a - c, 2a + b + 3c, a + b + 4c)$$

This is the required linear transformation.

Exercises:

1. Obtain the matrices for the following linear transformations.

(a) $T: V_2(\mathbb{R}) \rightarrow V_2(\mathbb{R})$ given by $T(a, b) = (-b, a)$ w.r.t.

(i) Standard basis

(ii) The basis $\{(1, 2), (1, -1)\}$ for both domain and range.

(b) $T: V_3(\mathbb{R}) \rightarrow V_2(\mathbb{R})$ given by $T(a, b, c) = (a + b, 2c - a)$ w.r.t.

(i) Standard basis

(ii) $\{(1, 0, -1), (1, 1, 1), (1, 0, 0)\}$ as a basis for $V_3(\mathbb{R})$ and $\{(0, 1), (1, 0)\}$ for $V_2(\mathbb{R})$.

(c) $T: V_3(\mathbb{R}) \rightarrow V_3(\mathbb{R})$ given by

$$T(a, b, c) = (3a + c, -2a + b, a + 2b + 4c) \text{ w.r.t.}$$

(i) The standard basis

(ii) The basis $\{(1, 0, 1), (-1, 2, 1), (2, 1, 1)\}$ for both domain and range.

(d) Let V be the set of all polynomials of degree $\leq n$ in $\mathbb{R}[x]$.

$$T: V \rightarrow V \text{ defined by } T(f) = \frac{df}{dx} \text{ w.r.t. the basis}$$

$$\{1, x, x^2, \dots, x^n\}.$$

2. Obtain the linear transformation determined by the following matrices

(a) $T: V_2(\mathbb{R}) \rightarrow V_2(\mathbb{R})$ given by $\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$ w.r.t. the standard basis.

(b) $T: V_3(\mathbb{R}) \rightarrow V_3(\mathbb{R})$ given by $\begin{pmatrix} a & b & c \\ b & c & a \\ c & a & b \end{pmatrix}$ w.r.t. the standard basis.

(c) $T: V_2(\mathbb{R}) \rightarrow V_3(\mathbb{R})$ given by $\begin{pmatrix} 2 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix}$ w.r.t. the standard basis.

Answers:

1.

(a) (i) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ (ii) $\begin{pmatrix} -1/3 & -5/3 \\ 2/3 & 1/3 \end{pmatrix}$

(b) (i) $\begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 2 \end{pmatrix}$ (ii) $\begin{pmatrix} -3 & 1 \\ 1 & 2 \\ -1 & 1 \end{pmatrix}$

(c) (i) $\begin{pmatrix} 3 & 2 & -1 \\ 0 & 1 & 2 \\ 1 & 0 & 4 \end{pmatrix}$ (ii) $\begin{pmatrix} 17/4 & -3/4 & -1/2 \\ 35/4 & 15/4 & -7/2 \\ 17/2 & -3/2 & 0 \end{pmatrix}$

(d) $\begin{pmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 2 & 0 & \dots & 0 & 0 \\ 0 & 0 & 3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & n & 0 \end{pmatrix}$

2.

(a) $T(a, b) = (a \cos \theta + b \sin \theta, -a \sin \theta + b \cos \theta)$

(b) $T(x, y, z) = (ax + by + cz, bx + cy + az, cx + ay + bz)$

(c) $T(a, b) = (2a + b, a + b, -a - b)$.

Definition: 7.5.1

Let $A = (a_{ij})$ and $B = (b_{ij})$ be two $m \times n$ matrices. We define the **sum** of these two matrices by $A + B = (a_{ij} + b_{ij})$.

Note that we have defined addition only for two matrices having the same number of rows and the same number of columns.

Definition: 7.5.2

Let $A = (a_{ij})$ be an arbitrary matrix over a field F . let $\alpha \in F$. We define $\alpha A = (\alpha a_{ij})$.

Theorem: 7.22

The set $M_{mn}(F)$ of all $m \times n$ matrices over the field F is a vector space of dimension mn over F under matrix addition and scalar multiplication defined above.

Proof:

Let $A = (a_{ij})$ and $B = (b_{ij})$ be two $m \times n$ matrices over the field F . The addition of $m \times n$ matrices is a binary operation which is both commutative and associative. The $m \times n$ matrix whose entries are 0 is the **identity matrix** and $(-a_{ij})$ is the **inverse matrix** of (a_{ij}) . Thus the set of all $m \times n$ matrices over the field F is an **abelian group** with respect to addition. The verification of the following **axioms** are straight forward.

- (a) $\alpha(A + B) = \alpha A + \alpha B$
- (b) $(\alpha + \beta)A = \alpha A + \beta A$
- (c) $(\alpha\beta)A = \alpha(\beta A)$
- (d) $1A = A$.

Hence the set of all $m \times n$ over F is a vector space over F .

Now, we shall prove that the dimension of this vector space is mn . Let E_{ij} be the matrix with entry 1 in the $(i, j)^{\text{th}}$ place and 0 in the other places. We have mn matrices of this form. Also any matrix $A = (a_{ij})$ can be written as $A = \sum a_{ij} E_{ij}$. Hence A is a linear combination of the matrices E_{ij} . Further these mn matrices E_{ij} are linearly independent. Hence these mn matrices form a basis for the space of all $m \times n$ matrices. Therefore the dimension of the vector space is mn .

Theorem: 7.23

Let V and W be two finite dimensional vector space over a field F . Let $\dim V = m$ and $\dim W = n$. Then $L(V, W)$ is a vector space of dimension mn over F .

Proof:

By theorem 6.8, $L(V, W)$ is a vector space over F . Now, we shall prove that the vector space over F . Now, we shall prove that the vector space $L(V, W)$ is isomorphic to the vector space $M_{mn}(F)$. Since $M_{mn}(F)$ is of dimension mn , it follows that $L(V, W)$ is also of dimension mn .

Fix a basis $\{v_1, v_2, \dots, v_m\}$ for V and a basis $\{w_1, w_2, \dots, w_n\}$ for W .

We know that any linear transformation.

$T \in L(V, W)$ can be represented by an $m \times n$ matrix over F .

Let T be represented by $M(T)$. This function

$M : L(V, W) \rightarrow M_{mn}(F)$ is clearly 1 - 1 and onto.

Let $T_1, T_2 \in L(V, W)$ and $M(T_1) = (a_{ij})$ and $M(T_2) = (b_{ij})$.

$$M(T_1) = (a_{ij}) \Rightarrow T_1(v_i) = \sum_{j=1}^n a_{ij} w_j$$

$$M(T_2) = (b_{ij}) \Rightarrow T_2(v_i) = \sum_{j=1}^n b_{ij} w_j$$

$$\therefore (T_1 + T_2)(v_i) = \sum_{j=1}^n (a_{ij} + b_{ij}) w_j$$

$$\begin{aligned} \therefore M(T_1 + T_2) &= (a_{ij} + b_{ij}) \\ &= (a_{ij}) + (b_{ij}) \\ &= M(T_1) + M(T_2). \end{aligned}$$

Similarly $M(\alpha T_1) = \alpha M(T_1)$.

Hence M is the required isomorphism from $L(V, W)$ to $M_{mn}(F)$.

INNER PRODUCT SPACES

8.0 Introduction:

Upto this point we have dealt with the algebraic properties of a vector space and these properties are consequences of the basic operations, namely, vector addition and scalar multiplication defined in the vector space. We know that in the usual three dimensional vector space $V_3(\mathbb{R})$ it is possible to talk about the length of a vector and angle between two vectors. These concepts of length and angle can be defined in terms of the usual “dot product” or “scalar product” of two vectors. The dot product of $u = (a_1, b_1, c_1)$ and $v = (a_2, b_2, c_2)$ is defined by

$$u \cdot v = a_1 a_2 + b_1 b_2 + c_1 c_2.$$

We note that the length of u is given by $\sqrt{u \cdot u}$ and the angle θ between u and v is determined by $\cos \theta = \frac{u \cdot v}{\sqrt{u \cdot u} \sqrt{v \cdot v}}$. Hence u and v are perpendicular or orthogonal iff $u \cdot v = 0$.

An inner product on a vector space is a generalization of the dot product and in terms of such an inner product we can define the length of a vector and angle between two vectors. Our study about angle will be restricted to the concept of perpendicularity of two vectors.

Throughout this section we shall deal only with vector spaces over the field F of real or complex numbers.

8.1 Definition and Examples

Definition: 8.1.1

Let V be a vector space over F . An inner product on V is a function which assigns to each ordered pair of vectors u, v in V a scalar in F denoted by $\langle u, v \rangle$ satisfying the following conditions.

$$(i) \quad \langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$$

$$(ii) \quad \langle \alpha u, v \rangle = \alpha \langle u, v \rangle$$

$$(iii) \quad \langle u, v \rangle = \overline{\langle v, u \rangle}, \text{ where } \overline{\langle v, u \rangle} \text{ is the complex conjugate of } \langle v, u \rangle.$$

$$(iv) \quad \langle u, u \rangle \geq 0 \text{ and } \langle u, u \rangle = 0 \text{ iff } u = 0.$$

A vector space with an inner product defined on it is called an **inner product space**. An inner product space is called an **Euclidean space** or **unitary space** according as F is the field of real numbers or complex numbers.

Note: 1

If F is the field of real numbers then condition (iii) takes the form $\langle u, v \rangle = \langle v, u \rangle$. Further (iii) asserts that $\langle u, u \rangle$ is always real and hence (iv) is meaningful whether F is the field of real or complex numbers.

Note: 2

$$\langle u, \alpha v \rangle = \bar{\alpha} \langle u, v \rangle.$$

$$\text{For, } \langle u, \alpha v \rangle = \overline{\langle \alpha v, u \rangle}$$

$$= \overline{\alpha \langle v, u \rangle}$$

$$= \bar{\alpha} \overline{\langle v, u \rangle}$$

$$= \bar{\alpha} \langle u, v \rangle.$$

Note: 3

$$\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$$

$$\text{For, } \langle u, v + w \rangle = \overline{\langle v + w, u \rangle}$$

$$= \overline{\langle v, u \rangle + \langle w, u \rangle}$$

$$= \overline{\langle v, u \rangle} + \overline{\langle w, u \rangle}$$

$$= \langle u, v \rangle + \langle u, w \rangle$$

Note: 4

$$\langle u, 0 \rangle = \langle 0, v \rangle = 0.$$

$$\text{For, } \langle u, 0 \rangle = \langle u, 0 \cdot u \rangle = 0 \langle u, u \rangle = 0.$$

$$\text{Similarly } \langle 0, v \rangle = 0.$$

Exercises:

Show that in an inner product space

$$(i) \quad \langle \alpha u + \beta v, w \rangle = \alpha \langle u, w \rangle + \beta \langle v, w \rangle$$

$$(ii) \quad \langle u, \alpha v + \beta w \rangle = \bar{\alpha} \langle u, v \rangle + \bar{\beta} \langle u, w \rangle$$

$$(iii) \quad \langle \alpha u + \beta v, \gamma w + \delta z \rangle = \alpha \bar{\gamma} \langle u, w \rangle + \alpha \bar{\delta} \langle u, z \rangle + \beta \bar{\gamma} \langle v, w \rangle + \beta \bar{\delta} \langle v, z \rangle$$

where $\alpha, \beta, \gamma, \delta \in F$ and $u, v, w, z \in V$.

Examples:

1. $V_n(\mathbb{R})$ is a real inner product space with inner product defined by

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2 + \dots + x_n y_n \text{ where}$$

$$x = (x_1, x_2, \dots, x_n) \text{ and}$$

$$y = (y_1, y_2, \dots, y_n).$$

This is called the standard inner product on $V_n(\mathbb{R})$

Proof:

Let $x, y, z \in V_n(\mathbb{R})$ and $\alpha \in \mathbb{R}$.

$$\begin{aligned} (i) \quad \langle x + y, z \rangle &= (x_1 + y_1)z_1 + (x_2 + y_2)z_2 + \dots + (x_n + y_n)z_n \\ &= (x_1 z_1 + x_2 z_2 + \dots + x_n z_n) + (y_1 z_1 + y_2 z_2 + \dots + y_n z_n) \\ &= \langle x, z \rangle + \langle y, z \rangle. \end{aligned}$$

$$\begin{aligned} (ii) \quad \langle \alpha x, y \rangle &= \alpha x_1 y_1 + \alpha x_2 y_2 + \dots + \alpha x_n y_n \\ &= \alpha (x_1 y_1 + x_2 y_2 + \dots + x_n y_n) \\ &= \alpha \langle x, y \rangle. \end{aligned}$$

$$\begin{aligned} (iii) \quad \langle x, y \rangle &= x_1 y_1 + x_2 y_2 + \dots + x_n y_n \\ &= y_1 x_1 + y_2 x_2 + \dots + y_n x_n \\ &= \langle y, x \rangle. \end{aligned}$$

$$\begin{aligned} (iv) \quad \langle x, x \rangle &= x_1^2 + x_2^2 + \dots + x_n^2 \geq 0 \text{ and } \langle x, x \rangle = 0 \text{ iff} \\ x_1 &= x_2 = \dots = x_n = 0 \\ \therefore \quad \langle x, x \rangle &= 0 \text{ iff } x = 0 \end{aligned}$$

2. $V_n(\mathbb{C})$ is a complex inner product space with inner product defined by

$$\langle x, y \rangle = x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n \text{ where}$$

$$x = (x_1, x_2, \dots, x_n) \text{ and}$$

$$y = (y_1, y_2, \dots, y_n).$$

Proof:Let $x, y, z \in V_n(C)$ and $\alpha \in C$.

$$\begin{aligned}
 \text{(i)} \quad \langle x + y, z \rangle &= (x_1 + y_1)\overline{z_1} + (x_2 + y_2)\overline{z_2} + \dots + (x_n + y_n)\overline{z_n} \\
 &= (x_1\overline{z_1} + x_2\overline{z_2} + \dots + x_n\overline{z_n}) + (y_1\overline{z_1} + y_2\overline{z_2} + \dots + y_n\overline{z_n}) \\
 &= \langle x, z \rangle + \langle y, z \rangle.
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad \langle \alpha x, y \rangle &= \alpha x_1\overline{y_1} + \alpha x_2\overline{y_2} + \dots + \alpha x_n\overline{y_n} \\
 &= \alpha(x_1\overline{y_1} + x_2\overline{y_2} + \dots + x_n\overline{y_n}) \\
 &= \alpha \langle x, y \rangle
 \end{aligned}$$

$$\begin{aligned}
 \text{(iii)} \quad \overline{\langle y, x \rangle} &= \overline{y_1\overline{x_1} + y_2\overline{x_2} + \dots + y_n\overline{x_n}} \\
 &= \overline{y_1}\overline{\overline{x_1}} + \overline{y_2}\overline{\overline{x_2}} + \dots + \overline{y_n}\overline{\overline{x_n}} \\
 &= x_1\overline{y_1} + x_2\overline{y_2} + \dots + x_n\overline{y_n} \\
 &= \langle x, y \rangle.
 \end{aligned}$$

$$\begin{aligned}
 \text{(iv)} \quad \langle x, x \rangle &= x_1\overline{x_1} + \dots + x_n\overline{x_n} \\
 &= |x_1|^2 + |x_2|^2 + \dots + |x_n|^2 \geq 0
 \end{aligned}$$

and $\langle x, x \rangle = 0$ iff $x = 0$.

3. Let V be the set of all continuous real valued functions defined on the closed interval $[0,1]$. V is a real inner product space with inner product defined by

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt.$$

Proof:Let $f, g, h \in V$ and $\alpha \in R$.

$$\begin{aligned}
 \text{(i)} \quad \langle f + g, h \rangle &= \int_0^1 (f(t) + g(t))h(t)dt \\
 &= \int_0^1 f(t)h(t)dt + \int_0^1 g(t)h(t)dt \\
 &= \langle f, h \rangle + \langle g, h \rangle.
 \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad \langle \alpha f, g \rangle &= \int_0^1 \alpha f(t)g(t)dt \\
 &= \alpha \int_0^1 f(t)g(t)dt
 \end{aligned}$$

$$= \alpha \langle f, g \rangle.$$

$$(iii) \quad \langle f, g \rangle = \int_0^1 f(t)g(t)dt$$

$$= \int_0^1 g(t)f(t)dt$$

$$= \langle g, f \rangle.$$

$$(iv) \quad \langle f, f \rangle = \int_0^1 [f(t)]^2 dt \geq 0 \text{ and } \langle f, f \rangle = 0 \text{ iff } f = 0$$

Exercises:

1. Show that $V_2(\mathbb{R})$ is an inner product space with inner product defined by

$$\langle x, y \rangle = x_1 y_1 + x_2 y_1 - x_1 y_2 + 4x_2 y_2$$

Where $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

2. Show that $V_2(\mathbb{C})$ is an inner product space with inner product defined by

$$\langle x, y \rangle = 2x_1 \overline{y_1} + x_1 \overline{y_2} + y_2 \overline{y_1} + x_2 \overline{y_2}$$

Where $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

3. Let V be the set of all continuous complex valued functions defined on the closed interval $[0,1]$. Show that V is a complex inner product

space with inner product defined by $\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt$.

4. Which of the following are inner products on $V_2(\mathbb{R})$ where $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

$$(a) \quad \langle x, y \rangle = x_1 y_1 + 2x_1 y_2 + 2x_2 y_1 + 5x_2 y_2.$$

$$(b) \quad \langle x, y \rangle = x_1^2 - 2x_1 y_2 - 2x_2 y_1 + y_1^2.$$

$$(c) \quad \langle x, y \rangle = 6x_1 y_1 + 7x_2 y_2.$$

$$(d) \quad \langle x, y \rangle = x_1 y_1 - 2x_2 y_1 - 2x_1 y_2 + 4x_2 y_2.$$

Answers:

4. (b) and (d) are not inner products.

Definition: 8.1.2

Let V be an inner product space and let $x \in V$. The **norm** or **length**

of x , denoted by $\|x\|$, is defined by $\|x\| = \sqrt{\langle x, x \rangle}$.

x is called **unit vector** if $\|x\| = 1$.

Problems:

Problem: 1

Let V be the vector space of polynomials with inner product given by $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$.

Let $f(t) = t + 2$ and $g(t) = t^2 - 2t - 3$.

Find (i) $\langle f, g \rangle$ (ii) $\|f\|$.

Solution:

$$\begin{aligned} \text{(i)} \quad \langle f, g \rangle &= \int_0^1 f(t)g(t)dt \\ &= \int_0^1 (t+2)(t^2-2t-3)dt \\ &= \int_0^1 (t^3-7t-6)dt \\ &= \left[\frac{t^4}{4} - \frac{7t^2}{2} - 6t \right]_0^1 \\ &= \frac{1}{4} - \frac{7}{2} - 6 \\ &= -\frac{37}{4} \end{aligned}$$

$$\begin{aligned} \text{(ii)} \quad \|f\|^2 &= \langle f, f \rangle \\ &= \int_0^1 [f(t)]^2 dt \\ &= \int_0^1 (t+2)^2 dt \\ &= \int_0^1 (t^2 + 4t + 4) dt \\ &= \left[\frac{t^3}{3} + 2t^2 + 4t \right]_0^1 \\ &= \frac{1}{3} + 2 + 4 \end{aligned}$$

$$= \frac{19}{3}$$

$$\therefore \|f\| = \frac{\sqrt{19}}{\sqrt{3}}$$

Exercises:

1. Find the norm of the following vectors in $V_3(\mathbb{R})$ with standard inner product.

(a) $(1, 1, 1)$

(b) $(1, 2, 3)$

(c) $(3, -4, 0)$

(d) $4x + 5y$ where $x = (1, -1, 0)$ and $y = (1, 2, 3)$.

2. Find the set of all unit vectors in $V_3(\mathbb{R})$ with standard norm.

Answers:

1. (a) $\sqrt{3}$ (b) $\sqrt{14}$ (c) 5 (d) $3\sqrt{38}$

2. All points on the unit sphere with centre $(0, 0, 0)$.

Theorem: 8.1

The norm defined in an inner product space V has the following properties

(i) $\|x\| \geq 0$ and $\|x\| = 0$ iff $x = 0$.

(ii) $\|\alpha x\| = |\alpha| \|x\|$.

(iii) $|\langle x, y \rangle| \leq \|x\| \|y\|$ (Schwartz's inequality).

(iv) $\|x + y\| \leq \|x\| + \|y\|$ (Triangle inequality).

Proof:

(i) $\|x\| = \sqrt{\langle x, x \rangle} \geq 0$ and $\|x\| = 0$ iff $x = 0$.

(ii) $\|\alpha x\|^2 = \langle \alpha x, \alpha x \rangle$

$$= \alpha \langle x, \alpha x \rangle$$

$$= \alpha \bar{\alpha} \langle x, x \rangle$$

$$= |\alpha|^2 \|x\|^2.$$

Hence $\|\alpha x\| = |\alpha| \|x\|$.

(iii) The inequality is trivially true when $x = 0$ or $y = 0$. Hence let $x \neq 0$

and $y \neq 0$.

Consider $z = y - \frac{\langle y, x \rangle}{\|x\|^2} x$.

Then $0 \leq \langle z, z \rangle$

$$\begin{aligned} &= \left\langle y - \frac{\langle y, x \rangle}{\|x\|^2} x, y - \frac{\langle y, x \rangle}{\|x\|^2} x \right\rangle \\ &= \langle y, y \rangle - \frac{\langle y, x \rangle}{\|x\|^2} \langle y, x \rangle - \frac{\langle y, x \rangle}{\|x\|^2} \langle x, y \rangle + \frac{\langle y, x \rangle \overline{\langle y, x \rangle}}{\|x\|^2 \|x\|^2} \langle x, x \rangle \\ &= \|y\|^2 - \frac{\langle y, x \rangle \overline{\langle y, x \rangle}}{\|x\|^2} - \frac{\langle y, x \rangle \langle x, y \rangle}{\|x\|^2} + \frac{\langle y, x \rangle \overline{\langle y, x \rangle}}{\|x\|^2} \\ &= \|y\|^2 - \frac{\langle x, y \rangle \overline{\langle x, y \rangle}}{\|x\|^2} \end{aligned}$$

$$\therefore 0 \leq \|x\|^2 \|y\|^2 - |\langle x, y \rangle|^2$$

$$\therefore |\langle x, y \rangle| \leq \|x\| \|y\|.$$

$$\begin{aligned} \text{(iv)} \quad \|x + y\|^2 &= \langle x + y, x + y \rangle \\ &= \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &= \|x\|^2 + \langle x, y \rangle + \overline{\langle x, y \rangle} + \|y\|^2 \\ &= \|x\|^2 + 2 \operatorname{Re} \langle x, y \rangle + \|y\|^2 \\ &\leq \|x\|^2 + 2 |\langle x, y \rangle| + \|y\|^2 \\ &\leq \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 \quad (\text{by (iii)}). \\ &\leq (\|x\| + \|y\|)^2 \\ \therefore \|x + y\| &\leq \|x\| + \|y\|. \end{aligned}$$

Exercises:

1. Applying Schwartz's inequality to $V_n(\mathbb{C})$ with standard inner product, show that

$$\left| \sum_{i=1}^n x_i \overline{y_i} \right| \leq \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^n |y_i|^2 \right)^{1/2}$$

2. Show that in any inner product space V

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$$

3. Show that in any inner product space

$$\|\alpha x + \beta y\|^2 = |\alpha|^2 \|x\|^2 + \alpha \bar{\beta} \langle x, y \rangle + \bar{\alpha} \beta \langle y, x \rangle + |\beta|^2 \|y\|^2$$

4. Show that if equality is valid in Schwartz's inequality or triangle inequality then x and y are linearly dependent. Is the converse true?

5.

(a) Show that in a real inner product space, if $\langle x, y \rangle = 0$, then

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2.$$

(b) Show that in a real inner product space, if

$$\|x\|^2 + \|y\|^2 = \|x + y\|^2,$$

$$\text{Then } \langle x, y \rangle = 0.$$

6. In an inner product space we define the distance between any two vectors x and y by $d(x, y) = \|x - y\|$. Show that

(a) $d(x, y) \geq 0$ and $d(x, y) = 0$ iff $x = y$.

(b) $d(x, y) = d(y, x)$

(c) $d(x, y) \leq d(x, z) + d(z, y)$.

8.2 Orthogonality

Definition: 8.2.1

Let V be an inner product space and let $x, y \in V$. x is said to be **orthogonal** to y if $\langle x, y \rangle = 0$.

Note: 1

$$x \text{ is orthogonal to } y \Rightarrow \langle x, y \rangle = 0$$

$$\Rightarrow \overline{\langle x, y \rangle} = \bar{0}$$

$$\Rightarrow \langle y, x \rangle = 0$$

$$\Rightarrow y \text{ is orthogonal to } x.$$

Thus x and y are orthogonal iff $\langle x, y \rangle = 0$.

Note: 2

$$x \text{ is orthogonal to } y \Rightarrow \alpha x \text{ is orthogonal to } y.$$

Note: 3

$$x_1 \text{ and } x_2 \text{ are orthogonal to } y \Rightarrow x_1 + x_2 \text{ is orthogonal to } y.$$

Note: 4

$\mathbf{0}$ is orthogonal to every vector in V and is the only vector with this property.

Definition: 8.2.2

Let V be an inner product space. A set S of vectors in V is said to be an **orthogonal set** if any two distinct vectors in S are orthogonal.

Definition: 8.2.3

S is said to be an **orthonormal set** if S is orthogonal and $\|x\| = 1$ for all $x \in S$.

Example:

The standard basis $\{e_1, e_2, \dots, e_n\}$ in R^n or C^n is an orthogonal set with respect to the standard inner product.

Theorem: 8.2

Let $S = \{v_1, v_2, \dots, v_n\}$ be an orthogonal set of non-zero vectors in an inner product space V . Then S is linearly independent.

Proof:

$$\text{Let } \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = 0$$

$$\text{Then } \langle \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n, v_1 \rangle = \langle 0, v_1 \rangle = 0.$$

$$\therefore \alpha_1 \langle v_1, v_1 \rangle + \alpha_2 \langle v_2, v_1 \rangle + \dots + \alpha_n \langle v_n, v_1 \rangle = 0.$$

$$\therefore \alpha_1 \langle v_1, v_1 \rangle = 0 \quad (\text{since } S \text{ is orthogonal})$$

$$\therefore \alpha_1 = 0 \quad (\text{since } v_1 \neq 0).$$

$$\text{Similarly } \alpha_2 = \alpha_3 = \dots = \alpha_n = 0.$$

Hence S is linearly independent.

Theorem: 8.3

Let $S = \{v_1, v_2, \dots, v_n\}$ be an orthogonal set of non-zero vectors in V . Let $v \in V$ and $v = \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$. Then

$$\alpha_k = \frac{\langle v, v_k \rangle}{\|v_k\|^2}.$$

Proof:

$$\langle v, v_k \rangle = \langle \alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n, v_k \rangle$$

$$\begin{aligned}
 &= \alpha_1 \langle v_1, v_k \rangle + \alpha_2 \langle v_2, v_k \rangle + \dots + \alpha_k \langle v_k, v_k \rangle + \dots + \alpha_n \langle v_n, v_k \rangle \\
 &= \alpha_k \langle v_k, v_k \rangle \quad (\text{since } S \text{ is orthogonal}) \\
 &= \alpha_k \|v_k\|^2 \\
 \therefore a_k &= \frac{\langle v, v_k \rangle}{\|v_k\|^2}.
 \end{aligned}$$

Theorem: 8.4

Every finite dimensional inner product space has an orthonormal basis.

Proof:

Let V be a finite dimensional inner product space. Let $\{v_1, v_2, \dots, v_n\}$ be a basis for V . From this basis we shall construct an orthonormal basis $\{w_1, w_2, \dots, w_n\}$ by means of a construction known as Gram – Schmidt orthogonalisation process.

First we take $w_1 = v_1$.

$$\text{Let } w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1.$$

We claim that $w_2 \neq 0$. For, if $w_2 = 0$ then v_2 is a scalar multiple of w_1 and hence of v_1 which is a contradiction since v_1, v_2 are linearly independent.

$$\begin{aligned}
 \text{Also, } \langle w_2, w_1 \rangle &= \left\langle v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1, w_1 \right\rangle \\
 &= \left\langle v_2 - \frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} v_1, v_1 \right\rangle \quad (\because w_1 = v_1) \\
 &= \langle v_2, v_1 \rangle - \frac{\langle v_2, v_1 \rangle}{\|v_1\|^2} \langle v_1, v_1 \rangle \\
 &= \langle v_2, v_1 \rangle - \langle v_2, v_1 \rangle \\
 &= 0.
 \end{aligned}$$

Now, suppose that we have constructed non – zero orthogonal vectors $w_1, w_2, \dots, w_k$

$$\text{Then put } w_{k+1} = v_{k+1} - \sum_{j=1}^k \frac{\langle v_{k+1}, w_j \rangle}{\|w_j\|^2} w_j$$

We claim that $w_{k+1} \neq 0$. For, if $w_{k+1} = 0$, then v_{k+1} is a linear combination of $w_1, w_2, \dots, w_k$ and hence is a linear combination of $v_1, v_2, \dots, v_k$ which is a contradiction since $v_1, v_2, \dots, v_{k+1}$ are linearly independent.

Also

$$\begin{aligned}\langle w_{k+1}, w_i \rangle &= \langle v_{k+1}, w_i \rangle - \sum_{j=1}^k \frac{\langle v_{k+1}, w_j \rangle}{\|w_j\|^2} \langle w_j, w_i \rangle \\ &= \langle v_{k+1}, w_i \rangle - \frac{\langle v_{k+1}, w_i \rangle}{\|w_i\|^2} \langle w_i, w_i \rangle \\ &= \langle v_{k+1}, w_i \rangle - \langle v_{k+1}, w_i \rangle \\ &= 0\end{aligned}$$

Thus, continuing in this way we ultimately obtain a non-zero orthogonal set $\{w_1, w_2, \dots, w_n\}$.

By Theorem 8.2 this set is linearly independent and hence a basis.

To obtain an orthonormal basis we replace each w_i by $\frac{w_i}{\|w_i\|}$.

Problems:

Problem: 1

Apply Gram – Schmidt process to construct an orthonormal basis for $V_3(\mathbb{R})$ with the standard inner product for the basis $\{v_1, v_2, v_3\}$ where $v_1 = (1, 0, 1)$; $v_2 = (1, 3, 1)$ and $v_3 = (3, 2, 1)$.

Solution:

Take $w_1 = v_1 = (1, 0, 1)$

Then $\|w_1\|^2 = \langle w_1, w_1 \rangle = 1^2 + 0^2 + 1^2 = 2$.

and $\langle w_1, v_2 \rangle = 1 + 0 + 1 = 2$.

Put $w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1$

$= (1, 3, 1) - (1, 0, 1)$

$= (0, 3, 0)$.

Space for Hints

$$\therefore \|w_2\|^2 = 9.$$

$$\text{Also, } \langle w_2, v_3 \rangle = 0 + 6 + 0 = 6 \text{ and } \langle w_1, v_3 \rangle = 3 + 0 + 1 = 4.$$

$$\text{Now, } w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\|w_2\|^2} w_2$$

$$= (3, 2, 1) - \frac{4}{2}(1, 0, 1) - \frac{6}{9}(0, 3, 0)$$

$$= (3, 2, 1) - 2(1, 0, 1) - \frac{2}{3}(0, 3, 0)$$

$$= (1, 0, -1).$$

$$\therefore \|w_3\|^2 = 2.$$

$$\therefore \text{The orthogonal basis is } \{(1, 0, 1), (0, 3, 0), (1, 0, -1)\}.$$

$$\text{Hence the orthonormal basis is } \left\{ \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right), (0, 1, 0), \left(\frac{1}{\sqrt{2}}, 0, \frac{-1}{\sqrt{2}} \right) \right\}.$$

Problem: 2

Let V be the set of all polynomials of degree ≤ 2 together with the zero polynomial. V is a real inner product space with inner product defined by $\langle f, g \rangle = \int_{-1}^1 f(x)g(x)dx$. Starting with the basis $\{1, x, x^2\}$, obtain an orthonormal basis for V .

Solution:

$$\text{Let } v_1 = 1; v_2 = x \text{ and } v_3 = x^2.$$

$$\text{Let } w_1 = v_1.$$

$$\text{Then } \|w_1\|^2 = \langle w_1, w_1 \rangle = \int_{-1}^1 1dx = 2.$$

$$\text{Hence } \|w_1\| = \sqrt{2}.$$

$$w_2 = v_2 - \frac{\langle v_2, w_1 \rangle}{\|w_1\|^2} w_1$$

$$= x - \frac{1}{2} \int_{-1}^1 xdx$$

$$= x.$$

$$\therefore \|w_2\|^2 = \langle w_2, w_2 \rangle = \int_{-1}^1 x^2 dx = \frac{2}{3}.$$

$$\text{Now, } w_3 = v_3 - \frac{\langle v_3, w_1 \rangle}{\|w_1\|^2} w_1 - \frac{\langle v_3, w_2 \rangle}{\|w_2\|^2} w_2$$

$$= x^2 - \frac{1}{2} \int_{-1}^1 x^2 dx - \left(\frac{3x}{2} \right) \int_{-1}^1 x^3 dx$$

$$= x^2 - \frac{1}{3}.$$

$$\therefore \|w_3\|^2 = \langle w_3, w_3 \rangle = \int_{-1}^1 \left(x^2 - \frac{1}{3} \right)^2 dx = \frac{8}{45}$$

Hence the orthogonal basis is $\left\{ 1, x, x^2 - \frac{1}{3} \right\}$.

$\therefore$ The required orthonormal basis is

$$\left\{ \frac{1}{\sqrt{2}}, \frac{\sqrt{3}}{2}x, \frac{\sqrt{10}}{4}(3x^2 - 1) \right\}.$$

Problem: 3

Find a vector of unit length which is orthogonal to $(1, 3, 4)$ in $V_3(\mathbb{R})$ with standard inner product.

Solution:

Let $x = (x_1, x_2, x_3)$ be any vector orthogonal to $(1, 3, 4)$. Then $x_1 + 3x_2 + 4x_3 = 0$. Any solution of this equation gives a vector orthogonal to $(1, 3, 4)$. For example $x = (1, 1, -1)$ is orthogonal to $(1, 3, 4)$.

Also $\|x\| = \sqrt{3}$. Hence a unit vector orthogonal to $(1, 2, 3)$ is given by $\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{\sqrt{3}} \right)$.

Note:

The set of all vectors orthogonal to $(1, 3, 4)$ are the points lying on the plane $x + 3y + 4z = 0$, which is a two dimensional subspace of $V_3(\mathbb{R})$.

Problem: 4

Find an orthogonal basis containing the vector $(1, 3, 4)$ for $V_3(\mathbb{R})$ with the standard inner product.

Solution:

$(1,1,-1)$ is a vector orthogonal to $(1,3,4)$ (refer problem 3 above).

Now, let $y = (y_1, y_2, y_3)$ be a vector orthogonal to both $(1,3,4)$ and $(1,1,-1)$.

$$\text{Then } y_1 + 3y_2 + 4y_3 = 0$$

$$y_1 + y_2 - y_3 = 0.$$

Any solution of this system of equations gives a vector orthogonal to $(1,3,4)$ and $(1,1,-1)$.

For example $(7,-5,2)$ is one such vector. (by cross multiplication method).

Hence $\{(1,3,4), (1,1,-1), (7,-5,2)\}$ is an orthogonal basis containing $(1,3,4)$.

Exercises:

1. Applying Gram – Schmidt process find the orthonormal basis of $V_3(\mathbb{R})$ with the standard inner product starting with the following bases.

(a) $(1,-1,0), (2,-1,-2), (1,-1,-2)$

(b) $(2,-1,0), (4,-1,0), (4,0,-1)$

(c) $(1,0,1), (1,0,-1), (0,3,4)$

2. Let V be the set of all polynomials of degree ≤ 2 over $\mathbb{R}$ with inner product defined by $\langle f, g \rangle = \int_0^1 f(x)g(x)dx$. Starting with the basis

$\{1, x, x^2\}$ obtain an orthonormal basis for V .

3. Obtain an orthogonal basis for $V_3(\mathbb{R})$ with standard inner product containing the vectors.

(a) $(1,1,-1)$ and $(1,0,1)$

(b) $(7,-1,1)$.

Answers:

1.

(a) $\{\sqrt{2}/2, -\sqrt{2}/2, 0\}, (\sqrt{2}/6, \sqrt{2}/6, -2\sqrt{2}/3), (-2/3, -2/3, -1/3)\}$

$$(b) \{(2\sqrt{5}/5, -\sqrt{5}/5, 0), (\sqrt{5}/5, 2\sqrt{5}/5, 0), (0, 0, -1)\}$$

$$(c) \{(1/\sqrt{2}, 0, 1/\sqrt{2}), (1/\sqrt{2}, 0, -1/\sqrt{2}), (0, 1, 0)\}$$

$$2. \{1, (2x - 1/\sqrt{3}, (5x^2 - 6x + 1)\sqrt{5})\}.$$

8.3 Orthogonal Complement

Definition: 8.3.1

Let V be an inner product space. Let S be a subset of V . The orthogonal complement of S , denoted by $S^\perp$, is the set of all vectors in V which are orthogonal to every vector of S .

$$(i.e) S^\perp = \{x / x \in V \text{ and } \langle x, u \rangle = 0 \text{ for all } u \in S\}.$$

Examples:

1. $V^\perp = \{0\}$ and $\{0\}^\perp = V$ since 0 is the only vector which is orthogonal to every vector.

2. Let $S = \{(x, 0, 0) / x \in \mathbb{R}\} \subseteq V_3(\mathbb{R})$ with standard inner product.

Then

$$S^\perp = \{(0, y, z) / y, z \in \mathbb{R}\}.$$

(i.e) The orthogonal complement of the x - axis is the yz plane.

Theorem: 8.5

If S is any subset of V then $S^\perp$ is a subspace of V .

Proof:

Clearly $0 \in S^\perp$ and hence $S^\perp \neq \Phi$.

Now, let $x, y \in S^\perp$ and $\alpha, \beta \in F$.

Then $\langle x, u \rangle = \langle y, u \rangle = 0$ for all $u \in S$:

$$\therefore \langle \alpha x + \beta y, u \rangle = \alpha \langle x, u \rangle + \beta \langle y, u \rangle = 0 \text{ for all } u \in S.$$

$$\therefore \alpha x + \beta y \in S^\perp. \text{ Hence } S^\perp \text{ is a subspace of } V.$$

Theorem: 8.6

Let V be a finite dimensional inner product space. Let W be a subspace of V . Then V is the direct sum of W and $W^\perp$ (i.e) $V = W \oplus W^\perp$.

Proof:

We shall prove that

(i) $W \cap W^\perp = \{0\}$, and

(ii) $W + W^\perp = V$.

(I) Let $v \in W \cap W^\perp$. Then $v \in W$ and $v \in W^\perp$. Now, $v \in W^\perp \Rightarrow v$ is orthogonal to every vector in W .

In particular, v is orthogonal to itself.

$$\therefore \langle v, v \rangle = 0 \text{ and hence } v = 0.$$

$$\text{Hence } W \cap W^\perp = \{0\}.$$

(II) Let $\{v_1, v_2, \dots, v_r\}$ be an orthonormal basis for W . Let $v \in V$.

Consider

$$v_0 = v - \langle v, v_1 \rangle v_1 - \langle v, v_2 \rangle v_2 - \dots - \langle v, v_r \rangle v_r$$

$$\therefore \langle v_0, v_i \rangle = \langle v, v_i \rangle - \langle v, v_i \rangle \langle v_i, v_i \rangle \text{ (since } \langle v_i, v_j \rangle = 0 \text{ if } i \neq j)$$

$$= \langle v, v_i \rangle - \langle v, v_i \rangle \text{ (since } \langle v_i, v_i \rangle = 1)$$

$$= 0$$

$\therefore v_0$ is orthogonal to each of $v_1, v_2, \dots, v_r$ and hence is orthogonal to every vector in W . Hence $v_0 \in W^\perp$ and

$$v = [\langle v, v_1 \rangle v_1 + \langle v, v_2 \rangle v_2 + \dots + \langle v, v_r \rangle v_r] + v_0 \in W + W^\perp$$

$$\therefore V = W \oplus W^\perp.$$

Hence the theorem.

Corollary:

$$\dim V = \dim W + \dim W^\perp.$$

Proof:

$$\dim V = \dim(W \oplus W^\perp) = \dim W + \dim W^\perp.$$

Theorem: 8.7

Let V be a finite dimensional inner product space. Let W be a subspace of V . Then $(W^\perp)^\perp = W$.

Proof:

Let $w \in W$. Then for any $u \in W^\perp$, $\langle w, u \rangle = 0$

Hence $w \in (W^\perp)^\perp$. Thus $W \subseteq (W^\perp)^\perp$ (1)

Now by Theorem 8.6, $V = W \oplus W^\perp$.

Also $V = W^\perp \oplus (W^\perp)^\perp$.

Hence $\dim W = \dim (W^\perp)^\perp$ (2)

From (1) and (2) we get $W = (W^\perp)^\perp$.

Problems:

Problem: 1

Let V be an inner product space and let S_1 and S_2 be subsets of V . Then $S_1 \subseteq S_2 \Rightarrow S_2^\perp \subseteq S_1^\perp$.

Solution:

Let $u \in S_2^\perp$

Then $\langle u, v \rangle = 0$ for all $v \in S_2$.

But $S_1 \subseteq S_2$. Hence $\langle u, v \rangle = 0$ for all $v \in S_1$.

Hence $u \in S_1^\perp$. Thus $S_2^\perp \subseteq S_1^\perp$.

Problem: 2

Let W_1 and W_2 be subspaces of a finite dimensional inner product space. Then

$$(i) \quad (W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp.$$

$$(ii) \quad (W_1 \cap W_2)^\perp = W_1^\perp + W_2^\perp.$$

Solution:

(i) We know that $W_1 \subseteq W_1 + W_2$.

$$\therefore (W_1 + W_2)^\perp \subseteq W_1^\perp \text{ (by the problem 1).}$$

Similarly, $(W_1 + W_2)^\perp \subseteq W_2^\perp$

$$\text{Hence } (W_1 + W_2)^\perp \subseteq W_1^\perp \cap W_2^\perp. \quad (1)$$

Now, let $w \in W_1^\perp \cap W_2^\perp$.

Then $w \in W_1^\perp$ and $w \in W_2^\perp$.

$$\therefore \langle w, u \rangle = 0 \text{ for all } u \in W_1 \text{ and } W_2.$$

Now, let $v \in W_1 + W_2$.

Then $v = v_1 + v_2$ where $v_1 \in W_1$ and $v_2 \in W_2$.

$$\begin{aligned}
 \therefore \langle w, u \rangle &= \langle w, v_1 + v_2 \rangle \\
 &= \langle w, v_1 \rangle + \langle w, v_2 \rangle \\
 &= 0 + 0 \text{ (since } v_1 \in W_1 \text{ and } v_2 \in W_2) \\
 &= 0
 \end{aligned}$$

Hence $w \in (W_1 + W_2)^\perp$.

$$\therefore W_1^\perp \cap W_2^\perp \subseteq (W_1 + W_2)^\perp \quad (2)$$

Form (1) and (2) we get $(W_1 + W_2)^\perp = W_1^\perp \cap W_2^\perp$.

(ii) Proof is similar to that of (i).

Exercises:

1. Let V be an inner product space. Let S be any subspace of V . Then show that

$$S^\perp = [L(S)]^\perp.$$

2. Find a basis for the orthogonal complement of the subspace spanned by $(2, 1, -2)$ in $V_3(\mathbb{R})$.

UNIT – 9

MATRICES

9.0 Introduction

In chapter 9 we have introduced $m \times n$ matrices and we have represented linear transformations by these matrices. In this chapter we shall develop the general theory of matrices. Throughout this chapter we deal with matrices whose entries are from the field F of real or complex numbers.

9.1 Algebra of Matrices

We have already seen that an $m \times n$ matrix A is an array of mn numbers a_{ij} where $1 \leq i \leq m, 1 \leq j \leq n$ arranged in m rows and n columns as follows:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

We shall denote this matrix by the symbol (a_{ij}) . If $m = n$, A is called a **square matrix** of order n .

Definition: 9.1.1

Two matrices $A = (a_{ij})$ and $B = (b_{ij})$ are said to be **equal** if A and B have the same number of rows and columns and the corresponding entries in the two matrices are same.

Addition of matrices:

We have already defined the addition of two $m \times n$ matrices $A = (a_{ij})$ and $B = (b_{ij})$ by $A + B = (a_{ij} + b_{ij})$.

We note that we can add two matrices iff they have the same number of rows and columns.

Example:

$$\text{If } A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 9 & 5 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 4 \\ 2 & 1 \\ -1 & 0 \end{pmatrix} \text{ then } A + B = \begin{pmatrix} 1 & 6 \\ 5 & 5 \\ 8 & 5 \end{pmatrix}$$

Remark:

The set of all $m \times n$ matrices is an abelian group under matrix addition. The $m \times n$ matrix with each entry 0 is the **zero matrix** and is denoted by **0** and the additive inverse of matrix $A = (a_{ij})$ is $(-a_{ij})$ and is denoted by $-A$.

If $A = (a_{ij})$ is any matrix and α is any number (real or complex) we have defined the matrix αA by

$$\alpha A = (\alpha a_{ij}).$$

The set of all $m \times n$ matrices over the field $\mathbf{R}$ under matrix addition and scalar multiplication defined above is a vector space. This result is true if $\mathbf{R}$ is replaced by $\mathbf{C}$ or by any field F .

We now proceed to define multiplication of matrices. We have already defined the multiplication of 2×2 matrices, which we generalise in the following definition.

Definition: 9.1.2

Let $A = (a_{ij})$ be an $m \times n$ matrix and $B = (b_{ij})$ be an $n \times p$ matrix.

We define the **product AB** as the $m \times p$ matrix (c_{ij}) where the ij^{th} entry c_{ij} is given by

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{in}b_{nj} = \sum_{k=1}^n a_{ik}b_{kj}.$$

Note: 1

The product AB of two matrices is defined only when the number of columns of A is equal to the number of rows of B .

The entry c_{ij} of the product AB is found by multiplying i^{th} row of A and the j^{th} column of B . To multiply a row and a column, we multiply the corresponding entries and add.

Examples:

$$1. \text{ Let } A = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 1 & 1 \\ 1 & 5 \\ 3 & 2 \\ 1 & 0 \end{pmatrix}. \text{ } A \text{ is a } 3 \times 4 \text{ matrix and}$$

B is a 4×2 matrix. Hence the product AB is a 3×2 matrix and

$$AB = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & 2 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 5 \\ 3 & 2 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 10 & 5 \\ 6 & 12 \\ 2 & 1 \end{pmatrix}$$

Note that in this example the product BA is not defined. Even if the product BA is defined, AB need not be equal to BA .

$$2. \text{ Let } A = \begin{pmatrix} 2 & 4 & 0 \\ 9 & 3 & 1 \\ 4 & 7 & 2 \end{pmatrix} \text{ and } I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \text{ Then } AI = IA = A.$$

(Verify)

3. Consider the square matrix of order n given by

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ \cdot & \cdot & \cdot & \dots & \cdot \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Let A be any $m \times n$ matrix. Then $I_n A = A$. Also if A is an $m \times n$ matrix, $A I_n = A$. If A is any $n \times n$ matrix, $A I_n = A$. If A is any $n \times n$ matrix, $A I_n = I_n A = A$. I_n is called the **identity matrix** of order n . We shall denote the identity matrix of any order by the symbol I .

Problems:

Problem: 1

Show that the matrix $A = \begin{pmatrix} 2 & -3 & 1 \\ 3 & 1 & 3 \\ -5 & 2 & -4 \end{pmatrix}$ satisfies the equation

$$A(A - I)(A + 2I) = 0.$$

Solution:

$$A - I = \begin{pmatrix} 2 & -3 & 1 \\ 3 & 1 & 3 \\ -5 & 2 & -4 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -3 & 1 \\ 3 & 0 & 3 \\ -5 & 2 & -5 \end{pmatrix}$$

$$A + 2I = \begin{pmatrix} 4 & -3 & 1 \\ 3 & 3 & 3 \\ -5 & 2 & -2 \end{pmatrix}$$

Now,

$$A(A - I)(A + 2I)$$

$$= \begin{pmatrix} 2 & -3 & 1 \\ 3 & 1 & 3 \\ -5 & 2 & -4 \end{pmatrix} \begin{pmatrix} 1 & -3 & 1 \\ 3 & 0 & 3 \\ -5 & 2 & -5 \end{pmatrix} \begin{pmatrix} 4 & -3 & 1 \\ 3 & 3 & 3 \\ -5 & 2 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} -12 & -4 & -12 \\ -9 & -3 & -9 \\ 21 & 7 & 21 \end{pmatrix} \begin{pmatrix} 4 & -3 & 1 \\ 3 & 3 & 3 \\ -5 & 2 & -2 \end{pmatrix}$$

$$= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$= 0$$

Hence $A(A - I)(A + 2I) = 0$.

Problem: 2

Prove that $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^n = \begin{pmatrix} \lambda^n & n\lambda^{n-1} \\ 0 & \lambda^n \end{pmatrix}$.

Solution:

We prove this result by induction on n .

When $n = 1$ result is obviously true.

Let us assume that the result is true for $n = k$.

$$\therefore \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^k = \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix}$$

$$\begin{aligned} \therefore \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}^k \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} &= \begin{pmatrix} \lambda^k & k\lambda^{k-1} \\ 0 & \lambda^k \end{pmatrix} \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \\ &= \begin{pmatrix} \lambda^{k+1} & \lambda^k + k\lambda^k \\ 0 & \lambda^{k+1} \end{pmatrix} \\ &= \begin{pmatrix} \lambda^{k+1} & (k+1)\lambda^k \\ 0 & \lambda^{k+1} \end{pmatrix} \end{aligned}$$

$\therefore$ The result is true for $n = k + 1$

Hence the result is true for all positive integers n .

Exercises:

- Write down six pairs of matrices A and B such that the product AB is defined and in each case compute the product AB .
- (a) Show that if A is an $m \times n$ matrix, then AB and BA are both defined iff B is an $n \times m$ matrix.
(b) Write down six pairs of matrices A and B such that both AB and BA are defined and compute the products AB and BA .
- If A and B are two matrices such that AB and $A + B$ are both defined, show that A, B are square matrices of the same order.

$$4. \text{ Let } A = \begin{pmatrix} 1 & -2 & 4 \\ -3 & 0 & 2 \\ 7 & 4 & 3 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 3 & -3 \\ 0 & 0 & 1 \end{pmatrix}.$$

Compute A, B^2, AB and BA .

$$5. \text{ If } A = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 2 \end{pmatrix} \text{ show that } A^2 - 4A - 5I = 0.$$

$$6. \text{ If } A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 2 & 1 \\ 2 & 0 & 3 \end{pmatrix} \text{ prove that } A^3 - 6A^2 + 7A + 2I = 0.$$

7. Prove that if $A = \begin{pmatrix} 3 & -4 \\ 1 & -1 \end{pmatrix}$, then $A^k = \begin{pmatrix} 1+2k & -4k \\ k & 1-2k \end{pmatrix}$ for any positive integer k .
8. Decide which of the following statements are true and which are false.
- (a) For any two matrices A and B , $A + B$ is defined.
 - (b) AB is defined $\Rightarrow BA$ is defined.
 - (c) For any matrix A , A^2 is defined.
 - (d) For any square matrix A , A^2 is defined.
 - (e) Matrix addition is commutative.
 - (f) Matrix addition is associative.
 - (g) Matrix multiplication is commutative.
 - (h) If A and B are 3×3 matrices then $(A + B)^2 = A^2 + 2AB + B^2$.
 - (i) If A and B are 3×3 matrices then $(A + B)(A - B) = A^2 - B^2$.
 - (j) (h) and (i) are true if $AB = BA$.

Answers:

8. (a) F (b) F (c) F (d) T (e) T (f) T (g) F (h) F (i) F (j) T

Theorem: 9.1

Let A be an $m \times n$ matrix, B an $n \times p$ matrix and C a $p \times q$ matrix. Then $A(BC) = (AB)C$.

Proof:

Let $A = (a_{ij})$, $B = (b_{ij})$ and $C = (c_{ij})$. Let us find the rs^{th} entry in $A(BC)$.

The r^{th} row in A is $a_{r1}, a_{r2}, \dots, a_{rn}$. The s^{th} column in BC consists of the elements $\sum b_{1j}c_{js}, \dots, \sum b_{nj}c_{js}$. Hence the rs^{th} entry in $A(BC)$ is

$$a_{r1} \sum b_{1j}c_{js} + \dots + a_{rn} \sum b_{nj}c_{js}$$

$$= \sum_{i=1}^n a_{ri} \sum_{j=1}^p b_{ij}c_{js} = \sum_{i=1}^n \sum_{j=1}^p a_{ri}b_{ij}c_{js}$$

Let us now find the rs^{th} entry in $(AB)C$.

The r^{th} row in AB is $\sum a_{ri}b_{i1}, \sum a_{ri}b_{i2}, \dots, \sum a_{ri}b_{ip}$.

The s^{th} column in C is $c_{1s}, c_{2s}, \dots, c_{ps}$.

Hence the rs^{th} entry in $(AB)C$ is

$$\left(\sum a_{ri}b_{i1}\right)c_{1s} + \left(\sum a_{ri}b_{i2}\right)c_{2s} + \dots + \left(\sum a_{ri}b_{ip}\right)c_{ps} = \sum_{i=1}^n \sum_{j=1}^p a_{ri}b_{ij}c_{js}$$

Thus $A(BC) = (AB)C$.

Theorem: 9.2

Let U, V, W be vector spaces of dimensions m, n and p respectively over a field F with respective bases $\{u_1, u_2, \dots, u_m\}, \{v_1, v_2, \dots, v_n\}$, and $\{w_1, w_2, \dots, w_p\}$. Let $T_1 : U \rightarrow V$ and $T_2 : V \rightarrow W$ be linear transformations and $M(T_1)$ and $M(T_2)$ their corresponding matrices with respect to these bases.

$$\text{Then } M(T_2 \circ T_1) = M(T_1)M(T_2).$$

Proof:

$M(T_1)$ is an $m \times n$ matrix and $M(T_2)$ is an $n \times p$ matrix.

Hence the product $M(T_1)M(T_2)$ is defined and is an $m \times p$ matrix.

$$\text{Let } M(T_1) = (a_{ij}) \text{ and } M(T_2) = (b_{jk}).$$

$$\text{Then, } T_1(u_i) = \sum_{j=1}^n a_{ij}v_j \text{ and } T_2(v_j) = \sum_{k=1}^p b_{jk}w_k.$$

$$\begin{aligned} \therefore (T_2 \circ T_1)(u_i) &= T_2\left(\sum_{j=1}^n a_{ij}v_j\right) \\ &= \sum_{j=1}^n a_{ij}T_2(v_j) \\ &= \sum_{j=1}^n a_{ij} \sum_{k=1}^p b_{jk}w_k \\ &= \sum_{j=1}^n \sum_{k=1}^p (a_{ij}b_{jk})(w_k) \end{aligned}$$

Thus $M(T_2 \circ T_1) = M(T_1)M(T_2)$.

Note: 1

Thus multiplication of two matrices is equivalent to the composition of their corresponding linear transformations in the reverse

order. Since composition of linear transformation is associative we get matrix multiplication is associative.

Note: 2

Let $M_n(F)$ denote the set of all square matrices of order n over the field F . Then matrix multiplication is an associative binary operation on $M_n(F)$. If $A, B, C \in M_n(F)$ the two distributive laws.

$$A(B + C) = AB + AC \text{ and } (A + B)C = AC + BC \text{ can be verified.}$$

Since $M_n(F)$ is already an abelian group under matrix addition we see that $M_n(F)$ is a ring.

Exercises:

1. Using $A = \begin{pmatrix} 1 & -1 & 1 \\ 5 & 0 & 1 \end{pmatrix}, B = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 1 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$ test the

associative law $A(BC) = (AB)C$ for matrix multiplication.

2. Compute $(2 \ 1 \ -1) \begin{pmatrix} 4 & -1 & 2 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 4 \\ 3 \end{pmatrix}$

3. Find for what values of x will $(x \ 4 \ 1) \begin{pmatrix} 2 & 1 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 4 \end{pmatrix} \begin{pmatrix} x \\ 4 \\ 1 \end{pmatrix} = 0$

4. Given that $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{pmatrix} A \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}$ find the matrix A .

Answers:

2. (3) $3.x = -2 \pm i\sqrt{6}$

4. $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 2 & 5/2 \\ -1 & -1/3 & -2/3 \end{pmatrix}$

Definition: 9.1.3

Let $A = (a_{ij})$ be an $m \times n$ matrix. Then the $n \times m$ matrix $B = (b_{ij})$ where $b_{ij} = a_{ji}$ is called the **transpose** of the matrix A and it is denoted by

A^T . Thus A^T is obtained from the matrix A by interchanging its rows and columns and the $(i, j)^{\text{th}}$ entry of $A^T = (j, i)^{\text{th}}$ entry of A .

$$\text{For example, if } A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 0 & 1 \\ 0 & 3 & 1 & 5 \end{pmatrix} \text{ then } A^T = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 3 \\ 3 & 0 & 1 \\ 4 & 1 & 5 \end{pmatrix}$$

Clearly if A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix.

Theorem: 9.3

Let A and B be two $m \times n$ matrices. Then

- (i) $(A^T)^T = A$.
- (ii) $(A + B)^T = A^T + B^T$.

Proof:

$$\begin{aligned} \text{(i) The } (i, j)^{\text{th}} \text{ entry of } (A^T)^T &= (j, i)^{\text{th}} \text{ entry of } A^T \\ &= (i, j)^{\text{th}} \text{ entry of } A. \end{aligned}$$

$$\therefore (A^T)^T = A$$

$$\begin{aligned} \text{(ii) The } (i, j)^{\text{th}} \text{ entry of } (A + B)^T &= (j, i)^{\text{th}} \text{ entry of } A + B \\ &= (j, i)^{\text{th}} \text{ entry of } A + (j, i)^{\text{th}} \text{ entry of } B. \\ &= (i, j)^{\text{th}} \text{ entry of } A^T + (i, j)^{\text{th}} \text{ entry of } B^T. \\ &= (i, j)^{\text{th}} \text{ entry of } (A^T + B^T). \end{aligned}$$

$$\therefore (A + B)^T = A^T + B^T.$$

Theorem: 9.4

Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Then

$$(AB)^T = B^T A^T.$$

Proof:

By hypothesis AB is defined and it is an $m \times p$ matrix. Hence

$(AB)^T$ is a $p \times m$ matrix.

Further B^T is a $p \times n$ matrix and A^T is an $n \times m$ matrix.

Hence, the product $B^T A^T$ is defined and it is a $p \times m$ matrix.

Now, let $A = (a_{ij})$, $B = (b_{ij})$ and $(AB) = (c_{ij})$.

The $(i, j)^{\text{th}}$ entry of $AB = c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$.

$\therefore$ The $(i, j)^{\text{th}}$ entry of $(AB)^T = c_{ji} = \sum_{k=1}^n a_{jk} b_{ki}$

Now the i^{th} row of B^T is the i^{th} column of B and it consists of the elements $b_{1i}, b_{2i}, \dots, b_{ni}$. Also the j^{th} column of A^T is the j^{th} row of A and it consists of the elements $a_{j1}, a_{j2}, \dots, a_{jn}$. Hence the $(i, j)^{\text{th}}$ entry of $B^T A^T = b_{1i} a_{j1} + b_{2i} a_{j2} + \dots + b_{ni} a_{jn}$.

$$= \sum_{k=1}^n b_{ki} a_{jk}$$

$$= (i, j)^{\text{th}} \text{ entry of } (AB)^T.$$

Hence $(AB)^T = B^T A^T$.

Definition: 9.1.4

Let $A = (a_{ij})$ be a matrix with entries from the field of complex numbers. The conjugate of A , denoted by $\bar{A}$, is defined by $\bar{A} = (\bar{a}_{ij})$.

$\bar{A}^T$ is called the conjugate transpose of the matrix A .

For example

$$\text{If } A = \begin{pmatrix} 2 & 2+i & -i \\ 1+i & -3 & 4+3i \end{pmatrix} \text{ then } \bar{A} = \begin{pmatrix} 2 & 2-i & i \\ 1-i & -3 & 4-3i \end{pmatrix}.$$

Theorem: 9.5

Let A and B be matrices with entries from C . Then

- (i) $\overline{(\bar{A})} = A$.
- (ii) $\overline{A+B} = \bar{A} + \bar{B}$.
- (iii) $\overline{kA} = \bar{k} \bar{A}$, where $k \in C$.
- (iv) $A = \bar{A} \Leftrightarrow$ all entries of A are real.
- (v) $\overline{AB} = \bar{A} \bar{B}$ provided AB is defined.
- (vi) $(\bar{A})^T = \overline{A^T}$.

Proof:(i) To prove: $\overline{\overline{A}} = A$.Let $A = (a_{ij})$

$$\overline{A} = (\overline{a_{ij}}) \quad [\because \text{The conjugate of } A \text{ is } \overline{A}]$$

$$\overline{\overline{A}} = (\overline{\overline{a_{ij}}})$$

$$= (a_{ij})$$

$$\therefore \overline{\overline{A}} = A$$

(ii) To prove: $\overline{A+B} = \overline{A} + \overline{B}$ Let $A = a_{ij}$ and $B = b_{ij}$ be matrices of type $m \times n$. Then $A+B = c_{ij}$ is a matrix of type $m \times n$, where $c_{ij} = a_{ij} + b_{ij}$

$$\therefore (i, j)^{\text{th}} \text{ entry of } \overline{A+B} = \overline{c_{ij}}$$

$$= \overline{a_{ij} + b_{ij}}$$

$$= \overline{a_{ij}} + \overline{b_{ij}}$$

$$= (i, j)^{\text{th}} \text{ entry of } \overline{A} + (i, j)^{\text{th}} \text{ entry of } \overline{B}.$$

$$\therefore \overline{A+B} = \overline{A} + \overline{B}$$

(iii) To prove: $\overline{kA} = \overline{k} \overline{A}$, where $k \in \mathbb{C}$ Let $A = a_{ij}$.

$$\therefore kA = b_{ij} \text{ where } b_{ij} = ka_{ij}$$

$$\therefore \overline{kA} = (\overline{b_{ij}}) \text{ where } \overline{b_{ij}} = \overline{ka_{ij}}$$

$$= \overline{k} \overline{a_{ij}}$$

$$= \overline{k} \overline{a_{ij}}$$

$$\therefore \overline{kA} = \overline{k} \overline{A}$$

(iv) To prove: $A = \overline{A} \Leftrightarrow$ all entries of A are real. $\Rightarrow$ Suppose $A = \overline{A}$.To prove: A is real.Now, $A = \overline{A}$. Then $a_{ij} = \overline{a_{ij}} \quad \forall i, j$

Space for Hints

If $a_{ij} = x + iy$ then $\overline{a_{ij}} = x - iy$

$$\therefore a_{ij} = \overline{\overline{a_{ij}}}$$

$$\Rightarrow x + iy = x - iy$$

$$\Rightarrow x + iy - x + iy = 0$$

$$\Rightarrow 2iy = 0$$

$$\Rightarrow y = 0$$

$\therefore y = 0$, ie, $a_{ij} = x = \text{a real number}$ or all the entries of A are real.

$\Leftarrow$

Suppose all the entries of A are real then $a_{ij} = \overline{a_{ij}} \quad \forall i, j$

$$\therefore A = \overline{A}$$

(v) To prove: $\overline{AB} = \overline{A} \overline{B}$ provided AB is defined.

Let $A = (a_{ij})$ and $B = (b_{ij})$ be matrices of types $m \times n$ and $n \times p$ respectively.

Then $AB = (a_{ij})$ is of type $m \times p$, where $c_{ij} = \sum_{k=1}^n a_{ik} b_{kj}$

$$\therefore (i, j)^{\text{th}} \text{ entry of } \overline{AB} = \overline{C_{ij}}$$

$$= \overline{a_{i1} b_{1j} + a_{i2} b_{2j} + \dots + a_{in} b_{nj}}$$

$$= \overline{a_{i1}} \overline{b_{1j}} + \overline{a_{i2}} \overline{b_{2j}} + \dots + \overline{a_{in}} \overline{b_{nj}}$$

$$= \sum_{k=1}^n \overline{a_{ik}} \overline{b_{kj}}$$

$$= (i, j)^{\text{th}} \text{ entry of } \overline{A} \overline{B}$$

$$\therefore \overline{AB} = \overline{A} \overline{B}$$

(vi) To prove: $(\overline{A})^T = \overline{A^T}$

$$\text{Let } A = (a_{ij})$$

$$\Rightarrow \overline{A} = (\overline{a_{ij}})$$

$$\Rightarrow (\overline{A})^T = (\overline{a_{ji}}) \quad (1)$$

Also

$$A = (a_{ij})$$

$$A^T = (a_{ji})$$

$$(\overline{A^T}) = (\overline{a_{ji}}) \quad (2)$$

From (1) and (2) we get

$$(\overline{A})^T = \overline{A^T}$$

Hence the proof.

Exercises:

1. Let $A = \begin{pmatrix} 3 & 4 & 6 \\ -1 & 7 & 2 \\ 4 & 3 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 & 2 \\ -2 & 0 & 0 \\ 3 & 4 & 1 \end{pmatrix}$. Find $A^T, B^T, (A+B)^T, (AB)^T$ and $B^T A^T$.

2. Let $A = \begin{pmatrix} 2i & 3+4i & 0 \\ 1+i & 1-i & i \\ 3 & 2i & 4 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 2 & 6 \\ -1 & 4 & 6 \\ 2 & 0 & 2 \end{pmatrix}$.

Find $\overline{A}, \overline{A+B}, \overline{AB}, \overline{A^T}, \overline{B^T}, \overline{A^T B}$ and $\overline{AB^T}$.

Types of Matrices:

Definition: 9.1.5

An $1 \times n$ matrix is called a **row matrix**. Thus a row matrix consists of 1 row and n columns.

It is of the form $(a_{11}, a_{12}, a_{13}, \dots, a_{1n})$.

Definition: 9.1.6

An $m \times 1$ matrix is called a **column matrix**. Thus a column matrix consist of

m rows and 1 column and it is of the form

$$\begin{pmatrix} a_{11} \\ a_{21} \\ \dots \\ a_{m1} \end{pmatrix}$$

Definition:

Let $A = (a_{ij})$ be a square matrix. Then the elements $a_{11}, a_{22}, \dots, a_{nn}$ are called the diagonal elements of A and the diagonal elements constitute what is known as the **principal diagonal** of the matrix A . A square matrix is called a **diagonal matrix** if all the entries which do not belong to the principal are zero. Hence in a diagonal matrix $a_{ij} = 0$ if $i \neq j$.

For example $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ is a diagonal matrix.

Definition: 9.1.7

A diagonal matrix in which all the entries of the principal diagonal are equal is called a **scalar matrix**.

For example $\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ is a scalar matrix.

Definition: 9.1.8

A square matrix (a_{ij}) is called an **upper triangular matrix** if all the entries above the principal diagonal are zero.

Hence $a_{ij} = 0$ whenever $i < j$ in an upper triangular matrix.

Definition: 9.1.9

A square matrix (a_{ij}) is called a **lower triangular matrix** if all the entries below the principal diagonal are zero.

Hence $a_{ij} = 0$ whenever $i > j$ in an lower triangular matrix.

For example, $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$ is lower triangular $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 0 & 2 & 3 & 0 \\ 2 & 3 & 2 & 4 \end{pmatrix}$ is upper

triangular.

Clearly a square matrix is a diagonal matrix iff it is both lower triangular and upper triangular.

Definition: 9.1.10

A square matrix $A = (a_{ij})$ is said to be **symmetric** if $a_{ij} = a_{ji}$ for all i, j .

Example:

$\begin{pmatrix} a & b \\ b & a \end{pmatrix}, \begin{pmatrix} a & h & g \\ h & b & f \\ g & f & c \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 0 & 0 & 5 \\ 3 & 0 & 6 & 7 \\ 4 & 5 & 7 & 8 \end{pmatrix}$ are symmetric matrices.

Theorem: 9.6

A square matrix A is symmetric iff $A = A^T$.

Proof:

Let A be a symmetric matrix.

Then the $(i, j)^{\text{th}}$ entry of A

$$= (j, i)^{\text{th}} \text{ entry of } A.$$

$$= (i, j)^{\text{th}} \text{ entry of } A^T.$$

Hence $A = A^T$.

Conversely let $A = A^T$.

Then $(i, j)^{\text{th}}$ entry of A

$$= (i, j)^{\text{th}} \text{ entry of } A^T$$

$$= (j, i)^{\text{th}} \text{ entry of } A.$$

Hence A is symmetric.

Theorem: 9.7

Let A be any square matrix. Then $A + A^T$ is symmetric.

Proof:

$$(A + A^T)^T = A^T + (A^T)^T$$

$$= A^T + A$$

$$= A + A^T.$$

Hence $A + A^T$ is symmetric.

Theorem: 9.8

Let A and B be symmetric matrices of order n . Then

$$(i) \quad A + B \text{ is symmetric.}$$

Space for Hints

- (ii) AB is symmetric iff $AB = BA$.
- (iii) $AB + BA$ is symmetric.
- (iv) If A is symmetric, then kA is symmetric where $k \in F$.

Proof:

$$\begin{aligned} \text{(i)} \quad (A+B)^T &= A^T + B^T \\ &= A + B \quad (\text{since } A \text{ and } B \text{ are symmetric}) \end{aligned}$$

$\therefore A + B$ is symmetric.

(ii) AB is symmetric

$$\Leftrightarrow (AB)^T = AB$$

$$\Leftrightarrow B^T A^T = AB \quad (\text{by Theorem 9.4})$$

$$\Leftrightarrow BA = AB.$$

$$\begin{aligned} \text{(iii)} \quad (AB+BA)^T &= (AB)^T + (BA)^T \\ &= B^T A^T + A^T B^T \\ &= BA + AB \quad (\text{since } A \text{ and } B \text{ are symmetric}) \\ &= AB + BA. \end{aligned}$$

$\therefore AB + BA$ is symmetric.

$$\text{(iv)} \quad (kA)^T = kA^T = kA \quad (\text{since } A \text{ is symmetric}).$$

$\therefore kA$ is symmetric.

Definition: 9.1.11

A square matrix $A = (a_{ij})$ is said to be skew symmetric if $a_{ij} = -a_{ji}$, for all i, j .

Note:

Let A be a skew symmetric matrix. Then $a_{ii} = -a_{ii}$. Hence $2a_{ii} = 0$ (ie) $a_{ii} = 0$, for all i . Thus in a skew symmetric matrix all the diagonal entries are zero.

$$\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix}, \begin{pmatrix} 0 & -2 & 1 \\ 2 & 0 & -3 \\ -1 & 3 & 0 \end{pmatrix} \text{ are examples of skew symmetric matrices.}$$

Theorem: 9.9

A square matrix A is skew symmetric matrix iff $A = -A^T$.

Proof:

Let $A = (a_{ij})$ be an $n \times n$ matrix.

Then $A^T = (b_{ij})$ is an $n \times n$ matrix, where $b_{ij} = a_{ji}$ for all i, j .

$\Rightarrow$

Suppose A is skew-symmetric.

Then $a_{ij} = -a_{ji}$ for all i, j (by definition)

$$= -b_{ij} \text{ for all } i, j.$$

$$A = -A^T.$$

$\Leftarrow$

Suppose $A = -A^T$.

Then $a_{ij} = -b_{ij}$ for all i, j .

$$= -a_{ji} \text{ for all } i, j.$$

$\therefore A$ is skew-symmetric.

Theorem: 9.10

Let A be any square matrix. Then $A - A^T$ is skew symmetric.

Proof:

$$(A - A^T)^T = A^T - (A^T)^T$$

$$= A^T - A$$

$$= -(A - A^T)$$

Hence $A - A^T$ is skew symmetric.

Theorem: 9.11

Any square matrix A can be expressed uniquely as the sum of a symmetric matrix and a skew symmetric matrix.

Proof:

Let A be any square matrix. Then $A + A^T$ is symmetric matrix ($\because$ by theorem 9.7)

$$\therefore \frac{1}{2}(A + A^T) \text{ is also a symmetric matrix.}$$

$$\text{Also, } \frac{1}{2}(A - A^T) \text{ is a skew symmetric matrix (by Theorem 9.10)}$$

$$\text{Now, } A = \frac{1}{2}(A + A^T) + \frac{1}{2}(A - A^T).$$

$\therefore$ A is the sum of a symmetric matrix and a skew symmetric matrix.

Now, to prove the uniqueness, let $A = R + S$ where S is a symmetric matrix and R is a skew symmetric matrix. We claim that $S = \frac{1}{2}(A + A^T)$ and

$$R = \frac{1}{2}(A - A^T).$$

$$A = S + R \quad (1)$$

$$\therefore A^T = (S + R)^T$$

$$= S^T + R^T$$

$$= S - R \quad (\text{since } S \text{ is symmetric and } R \text{ is skew symmetric})$$

$$\therefore A^T = S - R \dots\dots\dots$$

From (1) and (2) we get

$$S = \frac{1}{2}(A + A^T) \text{ and } R = \frac{1}{2}(A - A^T).$$

Theorem: 9.12

Let A and B be skew symmetric matrices of order n. Then

- (i) $A + B$ is skew symmetric.
- (ii) kA is skew symmetric, where $k \in F$.
- (iii) A^{2n} is a symmetric matrix and A^{2n+1} is a skew symmetric matrix where n is any positive integer.

Proof:

Let A, B be skew symmetric.

$$(i) (A + B)^T = A^T + B^T$$

$$= -A - B \quad (\text{by Theorem 9.9})$$

$$= -(A + B).$$

$$\therefore A + B \text{ is skew symmetric.}$$

$$(ii) (KA)^T = KA^T = K(-A) = -KA$$

$$\therefore KA \text{ is skew symmetric.}$$

$$(iii) \text{Let } m \text{ be any positive integer.}$$

$$\begin{aligned}
 \text{Then } (A^m)^T &= (AA \dots m \text{ times})^T \\
 &= A^T A^T \dots A^T \text{ (m times)} \\
 &= (-A)(-A) \dots (-A) \text{ (m times)} \\
 &\quad (\text{since } A^T = -A) \\
 &= (-1)^m A^m
 \end{aligned}$$

$$\therefore (A^m)^T = \begin{cases} A^m & \text{if } m \text{ is even} \\ -A^m & \text{if } m \text{ is odd} \end{cases}$$

$\therefore A^m$ is symmetric when m is even and skew symmetric when m is odd.

Definition: 9.1.12

A square matrix $A = (a_{ij})$ is said to be a **Hermitian matrix** if $a_{ij} = \bar{a}_{ji}$ for all i, j . A is said to be a **skew Hermitian matrix** iff $\bar{a}_{ij} = -a_{ji}$ for all i, j .

Example:

$$\begin{pmatrix} 1 & -1+2i & 3+4i \\ -1-2i & -2 & 3 \\ 3-4i & 3 & 2 \end{pmatrix} \text{ is a Hermitian matrix.}$$

$$\begin{pmatrix} 0 & -a+ib \\ a+ib & 0 \end{pmatrix}, \begin{pmatrix} ib & c+id \\ -c+id & ib \end{pmatrix} \text{ are skew Hermitian matrices.}$$

Note:

1. Any Hermitian matrix over $\mathbb{R}$ is a symmetric matrix and any skew Hermitian matrix over $\mathbb{R}$ is a skew symmetric matrix.
2. Let $A = (a_{ij})$ be a Hermitian matrix. Then $a_{ii} = \bar{a}_{ii}$ and hence a_{ii} is real for all i .
3. Let $A = (a_{ij})$ be a skew Hermitian matrix. Then $a_{ii} = -\bar{a}_{ii}$ and hence $a_{ii} = 0$ or purely imaginary for all i .

Theorem: 9.13

Let A be a square matrix.

- (i) A is Hermitian iff $A = \bar{A}^T$.
- (ii) A is skew Hermitian iff $A = -\bar{A}^T$.

Proof:

(i) Let $A = (a_{ij})$ be an $n \times n$ matrix with entries from the field of complex numbers.

$\Rightarrow$

Suppose A is Hermitian.

Then $a_{ij} = \bar{a}_{ji}$ for all i, j .

$$\begin{aligned}\therefore (i, j)^{\text{th}} \text{ entry of } A &= a_{ij} \\ &= \bar{a}_{ji}\end{aligned}$$

$$= (j, i)^{\text{th}} \text{ entry of } \bar{A}.$$

$$= (i, j)^{\text{th}} \text{ entry of } (\bar{A})^T.$$

$$\therefore A = \bar{A}^T$$

$\Leftarrow$

Suppose $A = \bar{A}^T$

Then $(i, j)^{\text{th}} \text{ entry of } A = (i, j)^{\text{th}} \text{ entry of } (\bar{A})^T.$

$$\text{ie) } a_{ij} = \bar{a}_{ji}$$

$\therefore A$ is Hermitian.

(ii) Let $A = (a_{ij})$ be an $n \times n$ matrix with entries from the field of complex numbers.

$\Rightarrow$

Suppose A is Skew-Hermitian.

Then $a_{ij} = -\bar{a}_{ji}$ for all i, j .

$$\begin{aligned}\therefore (i, j)^{\text{th}} \text{ entry of } A &= a_{ij} \\ &= -\bar{a}_{ji}\end{aligned}$$

$$= -(j, i)^{\text{th}} \text{ entry of } \bar{A}.$$

$$= -(i, j)^{\text{th}} \text{ entry of } (\bar{A})^T$$

$$\therefore A = -\bar{A}^T$$

$\Leftarrow$

Suppose $A = -\bar{A}^T$.

Then $(i, j)^{\text{th}} \text{ entry of } A$.

$$= -(i, j)^{\text{th}} \text{ entry of } (\bar{A})^T$$

$$\text{ie) } a_{ij} = -\bar{a}_{ji}$$

$\therefore A$ is skew – Hermitian.

Theorem: 9.14

Let A and B be square matrices of the same order. Then

- (i) A, B are Hermitian $\Rightarrow A + B$ is Hermitian.
- (ii) A, B are skew Hermitian $\Rightarrow A + B$ is skew Hermitian.
- (iii) A is Hermitian $\Rightarrow iA$ is skew Hermitian.
- (iv) A is skew Hermitian $\Rightarrow iA$ is Hermitian.
- (v) A is Hermitian and k is real $\Rightarrow kA$ is Hermitian.
- (vi) A is skew Hermitian and k is real $\Rightarrow kA$ is skew Hermitian.
- (vii) A, B are Hermitian $\Rightarrow AB + BA$ is Hermitian.
- (viii) A, B are Hermitian $\Rightarrow AB - BA$ is skew Hermitian.

Proof:

$$\begin{aligned} \text{(i)} \quad \overline{(A+B)}^T &= (\bar{A} + \bar{B})^T \\ &= \bar{A}^T + \bar{B}^T \\ &= A + B \quad (\text{since } A \text{ and } B \text{ are Hermitian}) \end{aligned}$$

$\therefore A + B$ is Hermitian.

$$\begin{aligned} \text{(iii)} \quad \overline{-(iA)}^T &= (-i\bar{A})^T \\ &= i\bar{A}^T \\ &= iA \quad (\text{since } A \text{ is Hermitian}) \end{aligned}$$

$\therefore iA$ is skew Hermitian.

$$\begin{aligned} \text{(vii)} \quad \overline{(AB+BA)}^T &= (\bar{A}\bar{B} + \bar{B}\bar{A})^T = (\bar{A}\bar{B} + \bar{B}\bar{A})^T \\ &= (\bar{A}\bar{B})^T + (\bar{B}\bar{A})^T \\ &= \bar{B}^T \bar{A}^T + \bar{A}^T \bar{B}^T \\ &= BA + AB \\ &= AB + BA \end{aligned}$$

$\therefore AB + BA$ is Hermitian.

(ii) Let A and B are Skew Hermitian

Space for Hints

ie, $A = -\bar{A}^T$ and $B = -\bar{B}^T$

$$\begin{aligned}\therefore (\overline{A+B})^T &= (\bar{A} + \bar{B})^T = \bar{A}^T + \bar{B}^T \\ &= -A - B \\ &= -(A+B)\end{aligned}$$

$$\Rightarrow (A+B) = -(\overline{A+B})^T$$

$\therefore A+B$ is skew Hermitian.

(iv) Let A be a skew Hermitian

$$\text{ie, } A = -\bar{A}^T$$

$$\begin{aligned}\text{Now } (iA)^T &= (-i\bar{A})^T = -i(\bar{A})^T \\ &= i(-\bar{A})^T \\ &= iA\end{aligned}$$

$\therefore iA$ is Hermitian.

(v) Let A be a Hermitian.

$$\text{ie, } A = \bar{A}^T$$

$$\begin{aligned}\text{Let } K \text{ be real. Then } (\overline{KA})^T &= (\bar{K} \bar{A})^T \\ &= (K \bar{A})^T\end{aligned}$$

$$(\because K \text{ is real, } \bar{K} = K)$$

$$= K(\bar{A})^T$$

$$= KA$$

$\therefore KA$ is Hermitian.

(vi) Let A be a Skew Hermitian.

$$\text{ie, } A = -\bar{A}^T$$

$$\begin{aligned}\text{Let } K \text{ be real. Then } (\overline{KA})^T &= (\bar{K} \bar{A})^T \\ &= K(\bar{A})^T\end{aligned}$$

$$[\because K \text{ is real, } \bar{K} = K]$$

$$= K(-A)$$

$$= -KA$$

$$\therefore KA = -(\overline{KA})^T$$

$\therefore KA$ is Skew Hermitian.

(viii) Let A and B are Hermitian.

$$\text{ie, } A = \bar{A}^T \text{ and } B = \bar{B}^T$$

$$\text{Now } (\overline{AB - BA})^T = (\overline{AB} - \overline{BA})^T$$

$$= (\overline{AB})^T - (\overline{BA})^T$$

$$= (\bar{A} \bar{B})^T - (\bar{B} \bar{A})^T$$

$$= \bar{B}^T \bar{A}^T - \bar{A}^T \bar{B}^T$$

$$= BA - AB$$

$$= - (AB - BA)$$

$$\therefore AB - BA = -(\overline{AB - BA})^T$$

$\therefore AB - BA$ is Skew Hermitian.

Theorem: 9.15

Let A be any square matrix.

Then

(i) $A + \bar{A}^T$ is Hermitian.

(ii) $A - \bar{A}^T$ is skew Hermitian.

Proof:

(i) Let $A + \bar{A}^T = B$.

$$\text{Then } \bar{B} = \bar{A} + A^T$$

$$\therefore \bar{B}^T = (\bar{A} + A^T)^T$$

$$= \bar{A}^T + A$$

$$= B$$

Hence $A + \bar{A}^T$ is Hermitian.

(ii) Let $A - \bar{A}^T = B$

$$\text{Then } \bar{B} = \bar{A} - A^T$$

$$\therefore \bar{B}^T = (\bar{A} - A^T)^T$$

$$= \bar{A}^T - A$$

$$= -(A - \bar{A}^T)$$

$$= -B$$

Space for Hints

$$\therefore B = -\overline{B}^T$$

$$\therefore A - \overline{A}^T \text{ is Skew Hermitian.}$$

Theorem: 9.16

Any square matrix A can be uniquely expressed as the sum of a Hermitian matrix and a skew Hermitian matrix.

Proof:

Let A be any square matrix. Then $A + \overline{A}^T$ is a Hermitian matrix.

($\because$ by theorem 9.15(i))

$$\therefore \frac{1}{2}(A + \overline{A}^T) \text{ is also a Hermitian matrix.}$$

Also, $\frac{1}{2}(A - \overline{A}^T)$ is a skew Hermitian matrix. ($\because$ by theorem 9.15(ii))

$$\text{Now } A = \frac{1}{2}(A + \overline{A}^T) + \frac{1}{2}(A - \overline{A}^T)$$

$\therefore A$ is the sum of a Hermitian matrix and a skew Hermitian matrix.

To prove: Uniqueness:

Let $A = R + S$, where R is Hermitian matrix and S is skew Hermitian matrix.

Claim:

$$R = \frac{1}{2}(A + \overline{A}^T) \text{ and } S = \frac{1}{2}(A - \overline{A}^T)$$

$$\text{Now, } A = R + S \quad (1)$$

$$\therefore \overline{A} = \overline{R + S}$$

$$\therefore \overline{A}^T = (\overline{R + S})^T = (\overline{R} + \overline{S})^T$$

$$= \overline{R}^T + \overline{S}^T$$

$$\overline{A}^T = R - S \quad (2)$$

$$[\because R \text{ is Hermitian ie. } R = \overline{R}^T, S \text{ is Skew Hermitian ie. } S = -\overline{S}^T]$$

From (1) + (2) $\Rightarrow$

$$(A + \overline{A}^T) = 2R$$

$$\Rightarrow R = \frac{1}{2}(A + \overline{A}^T)$$

$$\text{From (1) - (2)} \Rightarrow (A - \bar{A}^T) = 2S$$

$$\Rightarrow S = \frac{1}{2}(A - \bar{A}^T)$$

This shows that the representation of A as the sum of a Hermitian and a skew hermitian matrix is unique.

Definition: 9.1.13

A real square matrix A is said to be **orthogonal** if $AA^T = A^T A = I$.

Example:

$$A = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \text{ is an orthogonal matrix (verify).}$$

Theorem: 9.17

Let A and B be orthogonal matrices of the same order. Then

- (i) A^T is orthogonal.
- (ii) AB is orthogonal.

Proof:

- (i) $A^T(A^T)^T = A^T A = I$. (Since A is orthogonal).

Similarly we can prove $(A^T)^T A^T = I$.

$\therefore A^T$ is orthogonal.

- (ii) $(AB)(AB)^T = (AB)(B^T A^T)$
 $= A(BB^T)A^T$
 $= AIA^T$ (since B is orthogonal)
 $= AA^T$
 $= I$.

Similarly $(AB)^T(AB) = I$.

Hence AB is orthogonal.

Definition: 9.1.14

A square matrix A is said to be an **unitary matrix** if $A\bar{A}^T = \bar{A}^T A = I$.

For example $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ is unitary.

Note:

Any unitary matrix over $\mathbf{R}$ is an orthogonal matrix.

Theorem: 9.18

If A and B are unitary matrices of the same order, then AB is also an unitary matrix.

Proof:

Given that A and B are unitary matrices. ie, $A\bar{A}^T = \bar{A}^T A = I$ and $B\bar{B}^T = \bar{B}^T B = I$.

To prove: AB is an unitary matrix.

Now, $(AB)(\overline{AB})^T = (AB)(\bar{A} \bar{B})^T$

$$= (AB)(\bar{B}^T \bar{A}^T)$$

$$= A(B\bar{B}^T)\bar{A}^T$$

$$= A\bar{A}^T \quad (\because B\bar{B}^T = I)$$

$$= A\bar{A}^T$$

$$= I$$

Similarly $(\overline{AB})^T(AB) = I$

Hence AB is an unitary matrix.

Exercises:

1. Give examples of each of the following types of matrices; upper triangular matrix, lower triangular matrix, diagonal matrix, scalar matrix, symmetric matrix, Hermitian matrix, skew Hermitian matrix, orthogonal matrix and unitary matrix.
2. Give examples of matrices over the field of complex numbers which are
 - (i) Symmetric but not Hermitian.
 - (ii) Skew symmetric but not skew Hermitian.
3. Show that the product of two upper (lower) triangular matrices of the same order is again an upper (lower) triangular matrix.
4. Show that the product of two diagonal matrices of the same order is again a diagonal matrix.
5. Show that any two diagonal matrices of the same order commute.

6. For any square matrix A show that AA^T and $A^T A$ are symmetric.
7. Show that if A is symmetric then A^T is symmetric.
8. Show that if A is skew symmetric then A^2 is symmetric and A^3 is skew symmetric.
9. Show that if A and B are symmetric matrices of the same order then $AB - BA$ is skew symmetric.
10. Show that if A and B are skew symmetric matrices then AB is symmetric iff $AB = BA$.
11. Show that any Hermitian matrix A can be written as $A = B + iC$ where B is a real symmetric matrix and C is a real skew symmetric matrix. State and prove that a similar result for a skew Hermitian matrix.
12. Show that every square matrix A can be uniquely expressed as $A = B + iC$ where B and C are Hermitian.
13. A square matrix A is called an **idempotent matrix** if $A^2 = A$.

Show that $\begin{pmatrix} 2 & -3 & 5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{pmatrix}$ and $\begin{pmatrix} -1 & 3 & 5 \\ 1 & -3 & 5 \\ -1 & 3 & 5 \end{pmatrix}$ are idempotent matrices.

14. Show that if $AB = A$ and $BA = B$ then A and B are idempotent matrices.
15. Show that if A is an idempotent matrix, then $B = I - A$ is also an idempotent matrix and $AB = BA = 0$.
16. A square matrix A is said to be **nilpotent** if $A^n = 0$ for some

positive integer n . Show that $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 3 \\ 5 & 2 & 6 \\ -2 & -1 & -3 \end{pmatrix}$ are nilpotent.

17. A square matrix A is said to be **involutory** if $A^2 = I$.

Space for Hints

Show that $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ are involutory.

18. Show that a square matrix A is involutory iff $(I + A)(I - A) = 0$.

19. Show that $\frac{1}{2} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{pmatrix}$ is an orthogonal matrix.

20. If (ℓ_i, m_i, n_i) where $i = 1, 2, 3$ are the direction cosines of three mutually perpendicular lines referred to an orthogonal cartesian coordinate system, then show that $\begin{pmatrix} \ell_1 & m_1 & n_1 \\ \ell_2 & m_2 & n_2 \\ \ell_3 & m_3 & n_3 \end{pmatrix}$ is an orthogonal matrix.

21. Show that $\frac{1}{2} \begin{pmatrix} \ell + i & -1 + i \\ 1 + i & 1 - i \end{pmatrix}$ is a unitary matrix.

22. Determine which of the following statements are true and which are false.

Let A and B be square matrices of order n .

- A, B are symmetric $\Rightarrow AB$ is symmetric.
- A, B are skew symmetric $\Rightarrow AB$ is skew symmetric.
- A, B are upper triangular matrices $\Rightarrow AB$ is upper triangular.
- A, B are lower triangular matrices $\Rightarrow AB$ is lower triangular.
- A, B are diagonal matrices $\Rightarrow AB$ is a diagonal matrix.
- A, B are scalar matrices $\Rightarrow AB$ is a scalar matrix.
- Conjugate of a symmetric matrix is symmetric.
- Conjugate of a skew symmetric matrix is skew symmetric.
- Conjugate of a Hermitian matrix is Hermitian.
- Conjugate of a skew Hermitian matrix is skew Hermitian.
- Any real symmetric matrix is Hermitian.

Answers:

22. (a) F (b) F (c) T (d) T
 (e) T (f) T (g) T (h) T
 (i) T (j) T (k) T

The Inverse of a Matrix

A 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has an inverse iff $|A| = ad - bc \neq 0$

and the inverse of A is given by $\frac{1}{|A|} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$. Such matrices are

called **non-singular**. In this section we shall describe the method of finding the inverse of any non-singular matrix of order n .

Determinants:

We can associate with any $n \times n$ matrix $A = (a_{ij})$ over a field F

an element of F given by the determinant

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

Its value can be determined in the usual way and it is denoted by $|A|$

For example,

$$(i) \quad \text{If } A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ then } |A| = ad - bc.$$

$$(ii) \quad \text{If } A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 2 & 1 \end{pmatrix} \text{ then } |A| = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 2 & 1 \end{vmatrix} = 1.$$

Definition: 9.1.15

A square matrix A is said to be **singular** if $|A| = 0$.

A is called a **non-singular matrix** if $|A| \neq 0$.

Remark:

The rule for multiplying two matrices is same as the rule for multiplying two determinants.

Hence if A and B are two $n \times n$ matrices.

$$|AB| = |A| |B|.$$

Theorem: 9.19

The product of any two non-singular matrices is non-singular.

Proof:

Let A and B be two non-singular matrices of the same order. Then $|A| \neq 0$ and $|B| \neq 0$.

$$\therefore |AB| = |A||B| \neq 0.$$

Hence AB is non-singular.

Note:

Sum of two non-singular matrices need not be non-singular. For, if A is any non-singular matrix then $-A$ is also a non-singular matrix and $A + (-A)$ is the zero matrix which is obviously a singular matrix.

Definition: 9.1.16

Let $A = (a_{ij})$ be an $n \times n$ matrix. If we delete the row and the column containing the element a_{ij} we obtain a square matrix of order $n-1$ and the determinant of this square matrix is called the **minor** of the element a_{ij} and is denoted by M_{ij} .

The minor M_{ij} multiplied by $(-1)^{i+j}$ is called the **cofactor** of the element a_{ij} and is denoted by A_{ij} .

$$\therefore A_{ij} = (-1)^{i+j} M_{ij}.$$

Example:

$$\text{Let } A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Corresponding to the 9 elements a_{ij} , we get 9 minors of A. For example, the minor of a_{11} is

$$M_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} \text{ and the minor of } a_{23} \text{ is}$$

$$M_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}.$$

The cofactor of a_{11} is $A_{11} = (-1)^{1+1} M_{11} = M_{11}$.

The cofactor of a_{23} is $A_{23} = (-1)^{2+3} M_{23} = -M_{23}$.

Definition: 9.1.17**Space for Hints**

Let $A = (a_{ij})$ be a square matrix. Let A_{ij} denote the co-factor of a_{ij} . The transpose of the matrix (A_{ij}) is called the **adjoint** or **adjugate** of the matrix A and is denoted by $\text{adj } A$.

Thus the $(i, j)^{\text{th}}$ entry of $\text{adj } A$ is A_{ji} .

Note:

If A is a square matrix of order n then $\text{adj } A$ is also a square matrix of order n .

Example:

$$\text{Let } A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{pmatrix}$$

$$\text{Then } A_{11} = \begin{vmatrix} 1 & -1 \\ 1 & 3 \end{vmatrix} = 4.$$

$$A_{12} = -\begin{vmatrix} 3 & -1 \\ -2 & 3 \end{vmatrix} = -7.$$

Similarly other co-factors can be calculated and we get

$$\text{adj } A = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix} = \begin{pmatrix} 4 & 2 & -2 \\ -7 & 7 & 7 \\ 5 & -1 & 1 \end{pmatrix}$$

We notice that

$$\begin{aligned} A(\text{adj } A) &= \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{pmatrix} \begin{pmatrix} 4 & 2 & -2 \\ -7 & 7 & 7 \\ 5 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 14 & 0 & 0 \\ 0 & 14 & 0 \\ 0 & 0 & 14 \end{pmatrix} \\ &= (\text{adj } A)A. \text{ (verify)} \end{aligned}$$

Exercises:

1. Write down six square matrices A and calculate $\text{adj } A$, $A(\text{adj } A)$ and $(\text{adj } A)A$.
2. Prove that $\text{adj } A^T = (\text{adj } A)^T$.
3. If A is symmetric prove that $\text{adj } A$ is symmetric.

Theorem: 9.20

Let A be any square matrix of order n . Then $(\text{adj } A)A = A(\text{adj } A) = |A| I$ where I is the identity matrix of order n .

Proof:

The $(i, j)^{\text{th}}$ element of $(A(\text{adj } A))$

$$= \sum_{k=1}^n a_{ik} A_{jk}$$

$$= \begin{cases} 0 & \text{if } i \neq j \\ |A| & \text{if } i = j \end{cases}$$

$$\therefore A(\text{adj } A) = \begin{pmatrix} |A| & 0 & \dots & 0 \\ 0 & |A| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |A| \end{pmatrix}$$

$$= |A| I.$$

Similarly, $(\text{adj } A)A = |A| I$.

Hence $(\text{adj } A)A = A(\text{adj } A) = |A| I$.

Note:

Suppose $|A| \neq 0$. Now, consider the matrix $B = \frac{1}{|A|} \text{adj } A$.

$$\text{Then } AB = A \left(\frac{1}{|A|} (\text{adj } A) \right)$$

$$= \frac{1}{|A|} (A \text{ adj } A)$$

$$= \frac{1}{|A|} |A| I$$

$$= I.$$

Similarly $BA = I$. Thus $AB = BA = I$.

Definition: 9.1.18

Let A be a square matrix of order n . A is said to be **invertible** in there exists a square matrix B of order n such that $AB = BA = I$ and B is called the **inverse** of A and is denoted by A^{-1} .

Note:

The invertible matrices are precisely the units of the ring $M_n(F)$

Theorem: 9.21

A square matrix A of order n is non-singular iff A is invertible.

Proof:

Suppose A is invertible.

Then there exists a matrix B such that $AB = BA = I$.

Hence $|AB| = |I| = 1$.

$\therefore |A| |B| = 1$.

Hence $|A| \neq 0$ so that A is non-singular.

Conversely, let A be non-singular. Hence $|A| \neq 0$.

Now, consider the matrix $B = \frac{1}{|A|} \text{adj} A$.

Then $AB = BA = I$. (refer the note above).

$\therefore A$ is invertible and B is the inverse of A .

Problems:**Problem: 1**

Compute the inverse of the matrix $A = \begin{pmatrix} 2 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{pmatrix}$.

Solution:

$$|A| = \begin{vmatrix} 2 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{vmatrix} = -1.$$

Since $|A| \neq 0$, A is non-singular.

Hence A^{-1} exists and is given by $A^{-1} = \frac{\text{adj} A}{|A|}$.

Now, we find $\text{adj} A = \begin{pmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{pmatrix}$ where $A_{ij}, (i, j = 1, 2, 3)$ are

cofactors of a_{ij} .

$$A_{11} = \begin{vmatrix} 6 & -5 \\ -2 & 2 \end{vmatrix} = 2;$$

$$A_{12} = -\begin{vmatrix} -15 & -5 \\ 5 & 2 \end{vmatrix} = 5$$

$$A_{13} = \begin{vmatrix} -15 & 6 \\ 5 & -2 \end{vmatrix} = 0;$$

$$A_{21} = -\begin{vmatrix} -1 & 1 \\ -2 & 2 \end{vmatrix} = 0$$

$$A_{22} = \begin{vmatrix} 2 & 1 \\ 5 & 2 \end{vmatrix} = -1;$$

$$A_{23} = -\begin{vmatrix} 2 & -1 \\ 5 & -2 \end{vmatrix} = -1$$

$$A_{31} = \begin{vmatrix} -1 & 1 \\ 6 & -5 \end{vmatrix} = -1$$

$$A_{32} = -\begin{vmatrix} 2 & 1 \\ -15 & -5 \end{vmatrix} = -5$$

$$A_{33} = \begin{vmatrix} 2 & -1 \\ -15 & 6 \end{vmatrix} = -3.$$

$$\text{Hence } \text{adj}A = \begin{pmatrix} 2 & 0 & -1 \\ 5 & -1 & -5 \\ 0 & -1 & -3 \end{pmatrix}$$

$$\therefore A^{-1} = \frac{1}{-1} \begin{pmatrix} 2 & 0 & -1 \\ 5 & -1 & -5 \\ 0 & -1 & -3 \end{pmatrix} = \begin{pmatrix} -2 & 0 & 1 \\ -5 & 1 & 5 \\ 0 & 1 & 3 \end{pmatrix}$$

Problem: 2

$$\text{If } \omega = e^{2\pi i/3} \text{ find the inverse of the matrix } A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{pmatrix}.$$

Solution:

We note that $\omega^3 = 1$.

$$\therefore |A| = \begin{vmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{vmatrix} = 1[\omega^2 - \omega^4] - 1[\omega - \omega^2] + 1[\omega^2 - \omega]$$

$$= \omega^2 - \omega - \omega + \omega^2 + \omega^2 - \omega = 3(\omega^2 - \omega).$$

Since $|A| \neq 0$, A is non-singular. Hence A^{-1} exists and is given

$$\text{by } A^{-1} = \frac{\text{adj} A}{|A|}.$$

To find: $\text{adj} A$

$$\text{adj} A = \begin{bmatrix} A_{11} & A_{21} & A_{31} \\ A_{12} & A_{22} & A_{32} \\ A_{13} & A_{23} & A_{33} \end{bmatrix}$$

Where A_{ij} , ($i, j = 1, 2, 3$) are cofactors of a_{ij} .

$$A_{11} = \begin{vmatrix} 1 & \omega \\ 1 & \omega^2 \end{vmatrix} = \omega^2 - \omega$$

$$A_{12} = -\begin{vmatrix} 1 & \omega^2 \\ 1 & \omega \end{vmatrix} = -(\omega - \omega^2) = \omega^2 - \omega$$

$$A_{13} = \begin{vmatrix} 1 & \omega \\ 1 & \omega^2 \end{vmatrix} = \omega^2 - \omega$$

$$A_{21} = \begin{vmatrix} 1 & 1 \\ \omega^2 & \omega \end{vmatrix} = -(\omega - \omega^2) = \omega^2 - \omega$$

$$A_{22} = \begin{vmatrix} 1 & 1 \\ 1 & \omega \end{vmatrix} = \omega - 1$$

$$A_{23} = -\begin{vmatrix} 1 & 1 \\ 1 & \omega^2 \end{vmatrix} = -(\omega^2 - 1) = 1 - \omega^2$$

$$A_{31} = \begin{vmatrix} 1 & 1 \\ \omega & \omega^2 \end{vmatrix} = \omega^2 - \omega$$

$$A_{32} = -\begin{vmatrix} 1 & 1 \\ 1 & \omega^2 \end{vmatrix} = -(\omega^2 - 1) = 1 - \omega^2$$

$$A_{33} = \begin{vmatrix} 1 & 1 \\ 1 & \omega \end{vmatrix} = \omega - 1$$

$$\therefore \text{adj } A = \begin{pmatrix} \omega^2 - \omega & \omega^2 - \omega & \omega^2 - \omega \\ \omega^2 - \omega & \omega - 1 & 1 - \omega^2 \\ \omega^2 - \omega & 1 - \omega^2 & \omega - 1 \end{pmatrix} \quad \text{i} \quad (\text{verify})$$

$$\begin{aligned} \therefore A^{-1} &= \frac{1}{3(\omega^2 - \omega)} \begin{pmatrix} \omega^2 - \omega & \omega^2 - \omega & \omega^2 - \omega \\ \omega^2 - \omega & \omega - 1 & 1 - \omega^2 \\ \omega^2 - \omega & 1 - \omega^2 & \omega - 1 \end{pmatrix} \\ &= \frac{1}{3\omega} \begin{pmatrix} \omega & \omega & \omega \\ \omega & 1 & -1 - \omega \\ \omega & -1 - \omega & 1 \end{pmatrix} \end{aligned}$$

Problem: 3

Show that a square matrix A is orthogonal iff $A^{-1} = A^T$.

Solution:

Suppose A is orthogonal. Then $AA^T = I$.

$$\therefore |AA^T| = |I| = 1.$$

$$\therefore |A| |A^T| = 1.$$

$$\therefore |A| |A| = 1.$$

$$\therefore |A| \neq 0 \text{ and hence } A \text{ is non-singular.}$$

$$\therefore A^{-1} \text{ exists.}$$

$$\text{Now, } A^{-1}(AA^T) = A^{-1}I.$$

$$\therefore (A^{-1}A)A^T = A^{-1}.$$

$$\therefore IA^T = A^{-1}$$

$$\therefore A^T = A^{-1}.$$

Conversely, let $A^T = A^{-1}$.

$$\text{Then } AA^T = AA^{-1} = I. \text{ Similarly } A^T A = I.$$

Hence A is orthogonal.

Problem: 4

Show that a square matrix A is involutory iff $A = A^{-1}$.

Solution:

Suppose A is involutory. Then $A^2 = I$. Hence $|A^2| = 1$.

$$\therefore |A|^2 = |A| |A| = 1.$$

$\therefore |A| \neq 0$ and hence A is non-singular.

$\therefore A^{-1}$ exists.

Now, $A^{-1}(AA) = A^{-1}I$.

$\therefore (A^{-1}A)A = A^{-1}$.

$\therefore IA = A^{-1}$.

$\therefore A = A^{-1}$

Conversely, let $A = A^{-1}$.

Then $A^2 = AA = AA^{-1} = I$

$\therefore A$ is involutory.

Exercises:

1. Compute the inverse of each of the following matrices.

(a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & 4 \\ -2 & 2 & 1 \end{pmatrix}$

(c) $\begin{pmatrix} 2 & 2 & -3 \\ -3 & 2 & 2 \\ 2 & -3 & 2 \end{pmatrix}$

(d) $\begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$

2. Show that the set of all non-singular matrices of order n over a field F is a group under matrix multiplication.
3. If A and B are non-singular matrices of order n prove that $(AB)^{-1} = B^{-1}A^{-1}$.
4. If A is non-singular symmetric matrix prove that A^{-1} is also a symmetric matrix.
5. If A is a non-singular matrix, prove that $(A^T)^{-1} = (A^{-1})^T$.
6. If A is orthogonal, prove that A^{-1} is orthogonal.

7. Determine which of the following statements are true and which are false. Let A, B and C be square matrices of order n. Then

- (a) A, B are non-singular $\Rightarrow$ AB is non-singular.
- (b) A, B are non-singular $\Rightarrow$ A + B is non-singular.
- (c) A, B are singular $\Rightarrow$ AB is singular.
- (d) A is singular, B is non-singular $\Rightarrow$ AB is singular.
- (e) A is non-singular, B singular AB is singular.
- (f) AB is non-singular $\Rightarrow$ BA is non-singular.
- (g) AB = AC $\Rightarrow$ B = C.
- (h) AB = AC and A non-singular $\Rightarrow$ B = C.
- (i) AB = 0 $\Rightarrow$ A and B are singular.
- (j) A(B + C) = AB + AC.

Answers:

1.

$$(a) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

$$(b) -\frac{1}{31} \begin{pmatrix} -9 & 4 & 11 \\ -8 & 7 & -4 \\ -2 & -6 & -1 \end{pmatrix}$$

$$(c) \frac{1}{5} \begin{pmatrix} 2 & 1 & 2 \\ 2 & 2 & 1 \\ 1 & 2 & 2 \end{pmatrix}$$

$$(d) \begin{pmatrix} \cos \alpha & \sin \alpha & 0 \\ -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

7. (a) T (b) F (c) T (d) T (e) T (f) T (g) F (h) T (i) F
(j) T

In the following theorem we bring out the connection between non-singular linear transformations and non-singular matrices.

Theorem: 9.22

Let V and W be vector spaces of dimension n over a field F with bases $v_1, v_2, \dots, v_n$ and $w_1, w_2, \dots, w_n$ respectively. Then a linear

transformation $T : V \rightarrow W$ is non – singular iff the associated matrix is non – singular.

Proof:

Let $T : V \rightarrow W$ be a non – singular linear transformation.

Then T is 1 – 1 and onto.

Hence $T^{-1} : W \rightarrow V$ is also a linear transformation.

Let A and B be the matrices representing the linear transformation T and T^{-1} with respect to the chosen bases.

By theorem 9.2, multiplication of matrices A and B is equivalent to the composition of the corresponding linear transformation T and T^{-1}

Also $T \circ T^{-1}$ and $T^{-1} \circ T$ are identity transformations

Hence $AB = BA = I$. Thus A has an inverse B .

Hence A is non – singular.

Conversely, let A be a non – singular matrix. Then A^{-1} exists.

Let $S : W \rightarrow V$ be the linear transformation determined by the matrix A^{-1} .

It is easily verified that $T \circ S = S \circ T = I$

Hence T has an inverse linear transformation S .

Hence T is a non – singular linear transformation.

9.2 Elementary Transformations

Definition: 9.2.1

Let A be an $m \times n$ matrix over a field F . An elementary row – operation on A is of any one of the following three types.

1. The interchange of any two rows.
2. Multiplication of a row by a non – zero element c in F .
3. Addition of any multiple of one row with any other row.

Similarly we define an elementary column operation on A as any one of the following three types.

1. The interchange of any two columns.
2. Multiplication of a column by a non – zero element c in F .
3. Addition of any multiple of one column with any other column.

Example:

Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \\ 3 & -1 \end{pmatrix}$, $A_1 = \begin{pmatrix} 3 & -1 \\ 2 & 1 \\ 1 & 2 \end{pmatrix}$, $A_2 = \begin{pmatrix} 2 & 2 \\ 4 & 1 \\ 6 & -1 \end{pmatrix}$, $A_3 = \begin{pmatrix} 1 & 2 \\ 5 & 7 \\ 3 & -1 \end{pmatrix}$. A_1 is obtained

from A by interchanging the first and third rows.

A_2 is obtained from A by multiplying the first column of A by 2.

A_3 is obtained from A by adding to the second row the multiple by 3 of the first row.

Notation:

We shall employ the following notations for elementary transformations.

- (i) Interchanging of i^{th} and j^{th} rows will be denoted by $R_i \leftrightarrow R_j$.
- (ii) Multiplication of i^{th} row by a non - zero element $c \in F$ will be denoted by $R_i \rightarrow cR_i$.
- (iii) Addition of k times the j^{th} row to the i^{th} row will be denoted by $R_i \rightarrow R_i + kR_j$.

The corresponding column operations will be denoted by writing C in the place of R .

Definition: 9.2.2

An $m \times n$ matrix B is said to be **row equivalent** (**column equivalent**) to an $m \times n$ matrix A if B can be obtained from A by a finite succession of elementary row operations (column operations).

A and B are said to be **equivalent** if B can be obtained from A by a finite succession of elementary row or column operations.

If A and B are equivalent. We write $A \sim B$.

Exercise:

Prove that row equivalence, column equivalence and equivalence are equivalence relations in the set of all $m \times n$ matrices.

Definition: 9.2.3

A matrix obtained from the identity matrix by applying a single elementary row or column operation is called an elementary matrix.

For example, $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$ are elementary

matrices obtained from the identity matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ by applying the

elementary operations $R_1 \leftrightarrow R_2$, $R_1 \rightarrow 4R_1$, $R_3 \rightarrow R_3 + 2R_2$ respectively.

Exercise:

Give examples of elementary matrices of order 4.

Theorem: 9.23

Any elementary matrix is non – singular.

Proof:

The determinant of the identity matrix of any order is 1. Hence the determinant of an elementary matrix obtained by interchanging any two rows is -1. The determinant of an elementary matrix obtained by multiplying any row by $k \neq 0$ is k . The determinant of an elementary matrix obtained by adding a multiple of one row with another row is 1. Hence any elementary matrix is non – singular.

Theorem: 9.24

Let A be an $m \times n$ matrix and B be an $n \times p$ matrix. Then every elementary row (column) operation of the product AB can be obtained by subjecting the matrix A (matrix B) to the same elementary row (column) operation.

Proof:

Let $R_1, R_2, \dots, R_m$ denote the rows of the matrix A and $C_1, C_2, \dots, C_p$ denote the columns of B . By the definition of matrix multiplication

$$AB = \begin{pmatrix} R_1C_1 & R_1C_2 & \dots\dots\dots & R_1C_p \\ R_2C_1 & R_2C_2 & \dots\dots\dots & R_2C_p \\ \cdot & \cdot & \dots\dots\dots & \cdot \\ \cdot & \cdot & \dots\dots\dots & \cdot \\ \cdot & \cdot & \dots\dots\dots & \cdot \\ R_mC_1 & R_mC_2 & \dots\dots\dots & R_mC_p \end{pmatrix}$$

It is obvious from the above representation of AB that if we apply any elementary row operation on A the matrix AB is also subjected to the same elementary row operation. Also if we apply any elementary column operation on B the matrix AB is also subjected to the same elementary column operation.

Theorem: 9.25

Each elementary row operation on an $m \times n$ matrix A is equivalent to pre – multiplying the matrix A by the corresponding elementary $m \times m$ matrix.

Proof:

Since A is an $m \times n$ matrix we can write

$A = IA$ where I is the identity matrix of order m. By theorem 9.24 an elementary row operation IA is equivalent to the same row operation on I. But an elementary row operation on I gives an elementary matrix. Hence by pre – multiplying A by the corresponding elementary matrix we get the required row operation on A.

Note:

Similarly each elementary column operation of an $m \times n$ matrix A is equivalent to post – multiplying the matrix A by the corresponding elementary $n \times n$ matrix.

Corollary: 1

If two $m \times n$ matrices A and B are row equivalent then $A = PB$ where P is a non – singular $m \times m$ matrix.

Proof:

Since A is row equivalent to B, A can be obtained from B by applying successive elementary row operations. Hence $A = E_1E_2,\dots\dots\dots,E_nB$

where each E_i is an elementary matrix. Since each E_i is non-singular, $A = PB$ where $P = E_1 E_2, \dots, E_n$ and P is non-singular.

Corollary: 2

If two matrices A and B are column equivalent then $A = BQ$ where Q is non-singular matrix.

Corollary: 3

If two $m \times n$ matrices A and B are equivalent then $A = PBQ$ where P is a non-singular $m \times m$ matrix and Q is a non-singular $n \times n$ matrix.

Corollary: 4

The inverse of an elementary matrix is again an elementary matrix.

Proof:

Let E be an elementary matrix obtained from I by applying some elementary operations. If we apply the reverse operation on E , then E is carried back to I . Let E^* be the elementary matrix corresponding to the reverse operation.

Then $E^*E = EE^* = I$. Hence $E^* = E^{-1}$.

Hence E^{-1} is also an elementary matrix.

Canonical form of a matrix:

We now use elementary row and column operations to reduce any matrix to a simple form, called the canonical form of a matrix.

Theorem: 9.26

By successive applications of elementary row and column operations, any non-zero $m \times n$ matrix A can be reduced to a diagonal matrix D in which the diagonal entries are either 0 or 1 and all the 1's proceeding all the zeros on the diagonal. In other words, any non-zero

$m \times n$ matrix is equivalent to a matrix of the form $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ where I_r is

the $r \times r$ identity matrix and O is the zero matrix.

Proof:

We shall prove the theorem by induction on the number of rows

Space for Hints

of A . Suppose A has just one row. Let $A = (a_{11} a_{12}, \dots, a_{1n})$.

Since $A \neq 0$, by interchanging columns, if necessary, we can bring a non-zero entry c to the position a_{11} .

Multiplying A by c^{-1} we get 1 as the first entry.

Other entries in A can be made zero by adding suitable multiples of 1. Thus the result is true when $m=1$.

Now, suppose that the result is true for any non-zero matrix with $m-1$ rows.

Let A be a non-zero $m \times n$ matrix. By permuting rows and columns we can bring some non-zero entry c to the position a_{11} .

Multiplying the first row by c^{-1} we get 1 as the first entry.

All other entries in the first column can be made zero by adding suitable multiples of the first row to each other row.

Similarly all the other entries in the first row can be made zero.

This reduces A to a matrix of the form $B = \begin{pmatrix} I_1 & O \\ O & C \end{pmatrix}$ where C is an $(m-1) \times (n-1)$ matrix.

Now by induction hypothesis C can be reduced to the desired form by elementary row and column operations.

Hence A is equivalent to a matrix of the required form.

Corollary: 1

If A is an $m \times n$ matrix there exist non-singular square matrices P and Q of order m and n respectively such that $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$.

The result follows from corollary 3 of theorem 9.25.

Corollary: 2

Any non-singular square matrix A of order n is equivalent to the identity matrix.

Proof:

By corollary 1, $PAQ = \begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$.

Since P, A, Q are all non-singular $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$ is non-singular.

This is possible iff $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix} = I_n$.

Corollary: 3

Any non-singular matrix A can be expressed as a product of elementary matrices.

Proof:

By corollary 2, $PAQ = I_n$. Hence $A = P^{-1}Q^{-1}$. Further by corollary 4 of theorem 9.25, P^{-1} and Q^{-1} are products of elementary matrices.

Hence A is a product of elementary matrices.

Note:

The inverse of a non-singular matrix A can be computed by using elementary transformations. Let A be a non-singular matrix of order n . Then $AA^{-1} = A^{-1}A = I$. Now, the non-singular matrix A^{-1} can be expressed as the product of elementary matrices.

Let $A^{-1} = E_1E_2, \dots, E_n$.

Then $I = A^{-1}A = E_1E_2, \dots, E_nA$.

Thus every non-singular matrix A can be reduced to I by pre-multiplying A by elementary matrices.

Hence A can be reduced to the identity matrix by applying successive elementary row operations.

Now, $A = IA$. Reduce the matrix A in the left hand side to I by applying successive elementary row operations and apply the same elementary row operations and apply the same elementary row operations to the factor I in the right hand side.

Then we get $I = BA$ so that $B = A^{-1}$.

Problems:

Problem: 1

Reduce the matrix $A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 2 & 4 & -2 \end{pmatrix}$ to the canonical form.

Solution:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 2 & 4 & -2 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 2 & -1 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{array}{l} R_2 \rightarrow R_2 - R_1 \\ R_3 \rightarrow R_3 - 2R_1 \end{array}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{array}{l} C_2 \rightarrow C_2 - 2C_1 \\ C_3 \rightarrow C_3 + C_1 \end{array}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad C_3 \rightarrow C_3 + 3C_2$$

$$\sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad R_2 \rightarrow -R_2$$

Problem: 2

Find the inverse of the matrix $A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{pmatrix}$

Solution:

$$\begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & -1 \\ -2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot A$$

$$\Rightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -7 \\ 0 & 1 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \cdot A,$$

$$R_2 \rightarrow R_2 - 3R_1$$

$$R_3 \rightarrow R_3 + 2R_1$$

$$\Rightarrow \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -7 \\ 0 & 0 & 14 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 5 & -1 & 1 \end{pmatrix} \cdot A,$$

$$R_3 \rightarrow R_3 - R_2$$

$$\Rightarrow \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \frac{2}{7} & \frac{1}{7} & -\frac{1}{7} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{14} & -\frac{1}{14} & \frac{1}{14} \end{pmatrix}$$

$$R_1 \rightarrow R_1 - \frac{1}{7}R_3$$

$$R_2 \rightarrow R_2 + \frac{1}{2}R_3$$

$$R_3 \rightarrow \frac{1}{14}R_3$$

$$\Rightarrow A^{-1} = \begin{pmatrix} \frac{2}{7} & \frac{1}{7} & -\frac{1}{7} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{5}{14} & -\frac{1}{14} & \frac{1}{14} \end{pmatrix}$$

Exercises:

1. Write down six matrices (not necessarily square matrices) and reduced them to the canonical form.
2. Find the inverse of the following matrices by using elementary operations.

$$(a) \begin{pmatrix} 1 & -2 & 3 \\ 0 & -1 & 4 \\ -2 & 2 & 1 \end{pmatrix}$$

$$(b) \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

Answers:

$$2. (a) \begin{pmatrix} -9 & 8 & -5 \\ -8 & 7 & -4 \\ -2 & 2 & -1 \end{pmatrix}$$

$$(b) \frac{1}{2} \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}$$

Definition: 9.2.4

Let A and B be two square matrices of order n . B is said to be **similar** to A if there exists a $n \times n$ non-singular matrix P such that $B = P^{-1}AP$.

Problems:

Problem: 1

Similarity of matrices is an equivalence relation in the set of all $n \times n$ matrices.

Proof:

Let S be the set of all $n \times n$ matrices.

Let $A \in S$.

Since $A = I^{-1}AI$ and I is non-singular, A is similar to A .

Hence similarity of matrices is reflexive.

Now, let $A, B \in S$ and let A be similar to B .

$\therefore A = P^{-1}BP$ where $P \in S$ is a non-singular matrix.

Now, $P^{-1}BP = A \Rightarrow PP^{-1}BPP^{-1} = PAP^{-1}$

$$\Rightarrow B = PAP^{-1}$$

$$\Rightarrow B = (P^{-1})^{-1}A(P^{-1}).$$

Since P is non-singular $P^{-1} \in S$ is also non-singular.

$\therefore B$ is similar to A .

Hence similarity of matrices is symmetric.

Now, let $A, B, C \in S$.

Let A be similar to B to B be similar to C . Hence there exist non-singular matrices $P, Q \in S$ such that $A = P^{-1}BP$ and $B = Q^{-1}CQ$.

Now,

$$\begin{aligned} A &= P^{-1}BP \\ &= P^{-1}(Q^{-1}CQ)P \end{aligned}$$

$$= (P^{-1}Q^{-1})CQP$$

$$= (QP)^{-1}C(QP).$$

Since $P, Q \in S$ are non-singular, $QP \in S$ is also non-singular.

Hence A is similar to C .

$\therefore$ Similarity of matrices is transitive.

Hence similarity of matrices is an equivalence relation.

Problem: 2

If A and B are similar matrices show that their determinants are same.

Solution:

Let A and B be two similar matrices.

$\therefore$ there exists a non-singular matrix P such that

$$B = P^{-1}AP.$$

$$\text{Now, } |B| = |P^{-1}AP|$$

$$= |P^{-1}| |A| |P|$$

$$= |A| \quad \left(\text{since } |P^{-1}| = \frac{1}{|P|} \right)$$

Hence the result.

9.3 Rank of a Matrix

We now proceed to introduce the concept of the rank of a matrix.

Definition: 9.3.1

Let $A = (a_{ij})$ be an $m \times n$ matrix. The rows

$R_i = (a_{i1}, a_{i2}, \dots, a_{in})$ of A can be thought of as elements of F^n . The subspace of F^n generated by the m rows of A is called the **row space** of A .

Similarly, the subspace of F^m generated by the n columns of A is called the **column space** of A .

The dimension of the row space (column space) of A is called the **row rank** (**column rank**) of A .

Theorem: 9.27

Any two row equivalent matrices have the same row space and have the same row rank.

Proof:

Let A be an $m \times n$ matrix.

It is enough if we prove that the row space of A is not altered by any elementary row operation.

Obviously the row space of A is not altered by an elementary row operation of the type $R_i \leftrightarrow R_j$.

Now, consider the elementary row operation

$$R_i \rightarrow cR_i \text{ where } c \in F - \{0\}.$$

Since $L(\{R_1, R_2, \dots, R_i, \dots, R_n\}) = L(\{R_1, R_2, \dots, cR_i, \dots, R_n\})$
the row space of A is not altered by this type of elementary row operation.

Similarly we can easily prove that the row space of A is not altered by an elementary row operation of the type $R_i \rightarrow R_i + cR_j$.

Hence row equivalent matrices have the same row space and hence the same row rank.

Similarly we can prove the following theorem.

Theorem: 9.28

Any two column equivalent matrices have the same column rank.

Theorem: 9.29

The row rank and the column rank of any matrix are equal.

Proof:

Let $A = (a_{ij})$ be an $m \times n$ matrix.

Let $R_1, R_2, \dots, R_m$ denote the rows of A .

Hence $R_i = (a_{i1}, a_{i2}, \dots, a_{in})$

Suppose the row rank of A is r .

Then the dimension of the row space is r .

Let $v_1 = (b_{11}, \dots, b_{1n}), v_2 = (b_{21}, \dots, b_{2n}), \dots, v_r = (b_{r1}, \dots, b_{rn})$ be a basis for the row space of A.

Then each row is a linear combination of the vectors $v_1, v_2, \dots, v_r$.

Let, $R_1 = k_{11}v_1 + k_{12}v_2 + \dots + k_{1r}v_r$

$$R_2 = k_{21} v_1 + k_{22} v_2 + \dots + k_{2r} v_r$$

.....

$$R_m = k_{m1}v_1 + k_{m2} + \dots + k_{mr}v_r$$

Where $k_{ij} \in F$.

Equating the i^{th} component of each of the above equations, we get

$$a_{1i} = k_{11}b_{1i} + k_{12}b_{2i} + \dots + k_{1r}b_{ri}$$

$$a_{2i} = k_{21}b_{1i} + k_{22}b_{2i} + \dots + k_{2r}b_{ri}$$

$$\dots \dots \dots$$

$$\dots \dots \dots$$

$$a_{mi} = k_{m1}b_{1i} + k_{m2}b_{2i} + \dots + k_{mr}b_{ri}$$

Hence

$$\begin{pmatrix} a_{1i} \\ \cdot \\ \cdot \\ \cdot \\ a_{mi} \end{pmatrix} = b_{1i} \begin{pmatrix} k_{11} \\ \cdot \\ \cdot \\ \cdot \\ k_{m1} \end{pmatrix} + b_{2i} \begin{pmatrix} k_{12} \\ \cdot \\ \cdot \\ \cdot \\ k_{m2} \end{pmatrix} + \dots + b_{ri} \begin{pmatrix} k_{1r} \\ \cdot \\ \cdot \\ \cdot \\ k_{mr} \end{pmatrix}$$

Thus each column of A is a linear combination of r vectors.

Hence the dimension of the column space $\leq r$.

$\therefore$ Column rank of $A \leq r = \text{row rank of } A$.

Similarly, row rank of $A \leq \text{column rank of } A$.

Hence the row rank and the column rank of A are equal.

Definition: 9.3.2

The **rank** of a matrix A is the common value of its row and column rank.

Note: 1

Since the row rank and the column rank of a matrix are unaltered by elementary row and column operations, **equivalent matrices** have the same rank. In particular if a matrix A is reduced to its canonical form $\begin{pmatrix} I_r & O \\ O & O \end{pmatrix}$, then rank of $A = r$.

Thus to find the rank of a matrix A , we reduce A to the canonical form and find the number of non - zero entries in the diagonal.

Note that in the canonical form of the matrix A , there exists an $r \times r$ sub – matrix, namely, I_r , whose determinant is not zero.

Further every $(r+1) \times (r+1)$ sub – matrix contains a row of zeros and hence its determinant is zero.

Also under any elementary row or column operation the value of a determinant is either unaltered or multiplied by a non – zero constant.

Hence the matrix A is also such that

- (i) There exists an $r \times r$ sub – matrix whose determinant is non zero.
- (ii) The determinant of every $(r+1) \times (r+1)$ sub – matrix is zero.

Hence one can also define the rank of a matrix A to be r if A satisfies (i) and (ii).

Note: 2

Any non – singular matrix of order n is equivalent to the identity matrix and hence its rank is n .

Note: 3

The rank of a matrix is not altered on multiplication by non – singular matrices, since pre multiplication by non – singular matrices, since pre multiplication by a non – singular matrix is equivalent to applying elementary row operations and post – multiplication by a non – singular matrix is equivalent to applying elementary column operations.

Problems:

Problem: 1

Find the rank of the matrix $A = \begin{pmatrix} 4 & 2 & 1 & 3 \\ 6 & 3 & 4 & 7 \\ 2 & 1 & 0 & 7 \end{pmatrix}$

Solution:

$$A = \begin{pmatrix} 4 & 2 & 1 & 3 \\ 6 & 3 & 4 & 7 \\ 2 & 1 & 0 & 7 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 2 & 4 & 3 \\ 4 & 3 & 6 & 7 \\ 0 & 1 & 2 & 7 \end{pmatrix} C_1 \leftrightarrow C_3$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 4 & -5 & -10 & -5 \\ 0 & 1 & 2 & 7 \end{pmatrix} \begin{array}{l} C_1 \rightarrow C_2 - 2C_1 \\ C_3 \rightarrow C_3 - 4C_1 \\ C_4 \rightarrow C_4 - 3C_1 \end{array}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & -10 & -5 \\ 0 & 1 & 2 & 7 \end{pmatrix} R_1 \rightarrow R_2 - 4R_1$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 1 & 0 & 6 \end{pmatrix} \begin{array}{l} C_3 \rightarrow C_3 - 2C_2 \\ C_4 \rightarrow C_4 - C_2 \end{array}$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} R_3 \rightarrow R_3 + \frac{1}{5}R_2$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -5 & 0 & 0 \\ 0 & 0 & 6 & 0 \end{pmatrix} C_2 \leftrightarrow C_3$$

$$\sim \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{array}{l} R_2 \rightarrow -\frac{1}{5}R_2 \\ R_3 \rightarrow \frac{1}{6}R_3 \end{array}$$

$\therefore$ Rank of A = 3.

Problem: 2

Find the rank of the matrix $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 4 & 1 & 0 & 2 \\ 0 & 3 & 4 & 2 \end{pmatrix}$ by examining the

determinant minors.

Solution:

$$\begin{vmatrix} 1 & 1 & 1 \\ 4 & 1 & 0 \\ 0 & 3 & 4 \end{vmatrix} = 0 = \begin{vmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \\ 3 & 4 & 2 \end{vmatrix}$$

$$\begin{vmatrix} 1 & 1 & 1 \\ 4 & 1 & 2 \\ 0 & 3 & 2 \end{vmatrix} = 0 = \begin{vmatrix} 1 & 1 & 1 \\ 4 & 0 & 2 \\ 0 & 4 & 2 \end{vmatrix}$$

$\therefore$ Every 3 x 3 sub matrix of A has determinant zero.

$$\text{Also, } \begin{vmatrix} 1 & 1 \\ 4 & 1 \end{vmatrix} = -3 \neq 0$$

$\therefore$ Rank of A = 2.

Exercises:

1. Determine the rank of any six matrices of your choice.
2. Find the rank of the following matrices,

$$(a) \begin{pmatrix} 3 & -1 & 2 \\ 0 & 1 & -3 \\ 6 & -1 & 1 \end{pmatrix}$$

$$(b) \begin{pmatrix} 0 & 1 & 2 & 1 \\ 2 & -3 & 0 & -1 \\ 1 & 1 & -1 & 0 \end{pmatrix}$$

$$(c) \begin{pmatrix} 1 & -1 & 0 & 2 & 1 \\ 3 & 1 & 1 & -1 & 2 \\ 4 & 0 & 1 & 0 & 3 \\ 9 & -1 & 2 & 3 & 7 \end{pmatrix}$$

3. Find the column rank of the matrices

$$(a) \begin{pmatrix} 1 & 2 & -1 & 3 \\ 2 & 4 & 1 & -2 \\ 3 & 6 & 3 & -7 \end{pmatrix}$$

$$(b) \begin{pmatrix} 3 & 1 & -5 & -1 \\ 1 & -2 & 1 & -5 \\ 1 & 5 & -7 & 2 \end{pmatrix}$$

(Hint: Row rank = rank of the matrix = column rank)

4. Find the row rank of the matrix

$$\begin{pmatrix} 1 & 3 & 1 & -2 \\ 1 & 4 & 3 & -1 \\ 2 & 3 & -4 & -7 \\ 3 & 8 & 1 & -7 \end{pmatrix}$$

Answers:

$$2. (a) 2 \quad (b) 3 \quad (c) 3$$

$$3. (a) 2 \quad (b) 3$$

$$4. 2$$

9.4 Solving Linear Equations

In this section we shall apply the theory of matrices developed in the preceding sections to study the existence of solutions of simultaneous linear equations.

Matrix form of a set of linear equations

Consider a system of m linear equations in n unknowns $x_1, x_2, \dots, x_n$ given by

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots$$

$$\dots \quad \dots \quad \dots \quad \dots \quad \dots$$

$$a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m$$

Using the concept of matrix multiplication and equality of matrices this system can be written as $AX = B$ where,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

$$X = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix}, B = \begin{pmatrix} b_1 \\ b_2 \\ . \\ . \\ b_m \end{pmatrix}$$

The $m \times n$ matrix A is called the **coefficient matrix**.

Definition: 9.4.1

A set of values of $x_1, x_2, \dots, x_n$ which satisfy the above system of equations is called a **solution** of the system. The system of equations is said to be **consistent** if it has at least one solution. Otherwise the system is said to be **inconsistent**.

The $m \times (n+1)$ matrix given by

$$\begin{pmatrix} a_{11} & \dots & a_{1n} & b_1 \\ a_{21} & \dots & a_{2n} & b_2 \\ \vdots & \dots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{pmatrix}$$

is called the **augmented matrix** of the system and is denoted by (A, B) .

Thus the augmented matrix (A, B) is obtained by annexing to A to the column matrix B , which becomes the $(n+1)^{\text{th}}$ column in (A, B) .

Note:

Since every column in A appears in (A, B) the column space of the matrix A is a subspace of the column space of the matrix (A, B) .

Hence the rank of $A \leq$ rank of (A, B) .

Theorem: 9.30

The system of linear equations $AX = B$ is consistent iff rank of $A =$ rank of (A, B) .

Proof:

Let the system be consistent.

Let $u_1, u_2, \dots, u_n$ be a solution of the system.

Then $B = u_1 C_1 + u_2 C_2 + \dots + u_n C_n$ where $C_1, C_2, \dots, C_n$ denote the columns of A .

Hence the column space of the augmented matrix (A, B) , namely $\langle C_1, C_2, \dots, C_n, B \rangle$ is the same as the column space $\langle C_1, C_2, \dots, C_n \rangle$ of A .

Hence the rank of $A =$ rank of (A, B) .

Conversely let rank of $A =$ rank of (A, B) .

Then the column rank of $A =$ column rank of (A, B) .

$$\therefore \dim \langle C_1, C_2, \dots, C_n \rangle = \dim \langle C_1, C_2, \dots, C_n, B \rangle.$$

But $\langle C_1, C_2, \dots, C_n \rangle$ is a subspace of $\langle C_1, C_2, \dots, C_n, B \rangle$.

$\therefore B$ is a linear combination of $C_1, C_2, \dots, C_n$.

If $B = u_1 C_1 + \dots + u_n C_n$ then $u_1, u_2, \dots, u_n$ is a solution of the system.

Hence the theorem.

Remark:**Space for Hints**

The solution of a given system of simultaneous equations is not altered by interchanging any two equations or by multiplying any equation by a non – zero constant or by adding a multiple of one equation to another. Hence we can reduce the given system of equations to an equivalent system by applying elementary row operations to the augmented matrix. This reduced form will enable us to test for the consistency and to find the solution if it exists. This is illustrated in the following problems.

Problems:**Problem: 1**

Show that the equations

$$x + y + z = 6$$

$$x + 2y + 3z = 14$$

$$x + 4y + 7z = 30$$

are consistent and solve them.

Solution:

The given system of equations can be put in the matrix form

$$AX = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 7 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 14 \\ 30 \end{pmatrix} = B.$$

∴ The augmented matrix is given by

$$(A, B) = \begin{pmatrix} 1 & 1 & 1 & 6 \\ 1 & 2 & 3 & 14 \\ 1 & 4 & 7 & 30 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 3 & 6 & 24 \end{pmatrix} \begin{matrix} \\ R_2 \rightarrow R_2 - R_1 \\ R_3 \rightarrow R_3 - R_1 \end{matrix}$$

$$\sim \begin{pmatrix} 1 & 1 & 1 & 6 \\ 0 & 1 & 2 & 8 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} \\ \\ R_3 \rightarrow R_3 - 3R_2 \end{matrix}$$

Hence rank of A = rank of (A, B) = 2.

Hence the given system is consistent.

Also the given system of equations reduces to

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 8 \\ 0 \end{pmatrix}$$

$$\therefore x + y + z = 6$$

$$y + 2z = 8.$$

Putting $z = c$ we obtain the general solution of the system as
 $x = c - 2, y = 8 - 2c, z = c.$

Problem: 2

Verify whether the following system of equations is consistent. If it is consistent, find the solution.

$$x - 4y - 3z = -16$$

$$4x - y + 6z = 16$$

$$2x + 7y + 12z = 48$$

$$5x - 5y + 3z = 0.$$

Solution:

The matrix form of the system is given by

$$\begin{pmatrix} 1 & -4 & -3 \\ 4 & -1 & 6 \\ 2 & 7 & 12 \\ 5 & -5 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -16 \\ 16 \\ 48 \\ 0 \end{pmatrix}$$

$\therefore$ The augmented matrix is given by

$$(A, B) = \begin{pmatrix} 1 & -4 & -3 & -16 \\ 4 & -1 & 6 & 16 \\ 2 & 7 & 12 & 48 \\ 5 & -5 & 3 & 0 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & -4 & -3 & -16 \\ 0 & 15 & 18 & 80 \\ 0 & 15 & 18 & 80 \\ 0 & 15 & 18 & 80 \end{pmatrix} \begin{matrix} \\ R_2 \rightarrow R_2 - 4R_1 \\ R_3 \rightarrow R_3 - 2R_1 \\ R_4 \rightarrow R_4 - 5R_1 \end{matrix}$$

$$\sim \begin{pmatrix} 1 & -4 & -3 & -16 \\ 0 & 15 & 18 & 80 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{matrix} R_3 \rightarrow R_3 - R_2 \\ R_4 \rightarrow R_4 - R_2 \end{matrix}$$

$\therefore$ Rank of A = Rank of (A, B) = 2 and hence the system is consistent.

Also the system of equations reduces to

$$\begin{pmatrix} 1 & -4 & -3 \\ 0 & 15 & 18 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -16 \\ 80 \\ 0 \\ 0 \end{pmatrix}$$

$\therefore x - 4y - 3z = -16$ and $15y + 18z = 80$.

Putting $z = c$ we obtain the general solution of the systems as

$$x = -(9c/5) + (16/3),$$

$$y = -(6c/5) + (16/3);$$

$$z = c.$$

Problem: 3

For what values of η the equations

$$x + y + z = 1$$

$$x + 2y + 4z = \eta$$

$$x + 4y + 10z = \eta^2 \text{ are consistent?}$$

Solution:

The matrix form of the system is given by

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 10 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ \eta \\ \eta^2 \end{pmatrix}$$

$\therefore$ The augmented matrix is given by

$$(A, B) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & \eta \\ 1 & 4 & 10 & \eta^2 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \eta - 1 \\ 0 & 3 & 9 & \eta^2 - 1 \end{pmatrix} \begin{matrix} R_2 \rightarrow R_2 - R_1 \\ R_3 \rightarrow R_3 - R_1 \end{matrix}$$

Space for Hints

$$\sim \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & \eta - 1 \\ 0 & 0 & 0 & \eta^2 - 3\eta + 2 \end{pmatrix} R_3 \rightarrow R_3 - 3R_2$$

$\therefore$ The given system is consistent iff $\eta^2 - 3\eta + 2 = 0$.

$$\therefore \eta = 2 \text{ or } 1.$$

Problem: 4

Show that the system of equations

$$x + 2y + z = 11$$

$$4x + 6y + 5z = 8$$

$$2x + 2y + 3z = 19 \text{ is inconsistent.}$$

Solution:

The matrix form of the system is given by

$$\begin{pmatrix} 1 & 2 & 1 \\ 4 & 6 & 5 \\ 2 & 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 11 \\ 8 \\ 19 \end{pmatrix}$$

$\therefore$ The augmented matrix is given by

$$(A, B) = \begin{pmatrix} 1 & 2 & 1 & 11 \\ 4 & 6 & 5 & 8 \\ 2 & 2 & 3 & 19 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 2 & 1 & 11 \\ 0 & -2 & 1 & -36 \\ 0 & -2 & 1 & -3 \end{pmatrix} \begin{matrix} R_2 \rightarrow R_2 - 4R_1 \\ R_3 \rightarrow R_3 - 2R_1 \end{matrix}$$

$$\sim \begin{pmatrix} 1 & 2 & 1 & 11 \\ 0 & -2 & 1 & -36 \\ 0 & 0 & 0 & 33 \end{pmatrix} R_3 \rightarrow -R_3 - R_2$$

$\therefore$ Rank of $A = 2$ and rank of $(A, B) = 3$.

$\therefore$ The given system is inconsistent.

Exercises:

1. Solve or prove the inconsistency of the following systems of equations.

$$(a) \quad 2x - y + 3z = 8$$

$$x - 2y - z = -4$$

$$3x + y - 4z = 0$$

(b) $x + 2y - 5z = 0$

$$3x + 4y + 6z = 0$$

$$x + y + z = 0$$

(c) $x + 2y - 5z = -9$

$$3x - y + 2z = 5$$

$$2x + 3y - z = 3$$

$$4x - 5y + z = -3$$

(d) $x + y + z = 1$

$$x + 2y + 3z = 1$$

$$x + 3y + 5z = 7$$

$$x + 4y + 7z = 10$$

(e) $x - 2y - z - t = -1$

$$3x - 2z + 3t = -4$$

$$5x - 4y + t = -3$$

(f) $x + y + z = 7$

$$x + 2y + 3z = 8$$

$$y + 2z = 6$$

2. For what values of λ and μ the system of equations

$$x + y + z = 6$$

$$x + 2y + 3z = 10$$

$$x + 2y + \lambda z = \mu$$

is (a) inconsistent (b) consistent (c) consistent and the solution is unique.

3. Show that the set of all solutions of the system of homogeneous equations $AX = 0$ forms a vector space.

4. Show that a system of n equations in n unknowns given by $AX = Y$ has a unique solution if the $n \times n$ matrix A is non-singular.

Answers:

1.

(a) Consistent; $x = y = z = 2$

(b) Consistent; $x = y = z = 0$;

(c) Consistent; $x = \frac{1}{2}, y = \frac{3}{2}, z = \frac{5}{2}$;

(d) Consistent; $x = c - 2, y = 3 - 2c, z = c$

(e) Inconsistent;

(f) Inconsistent;

(g) Inconsistent.

2. If $\lambda = 3$ and $\mu \neq 10$, inconsistent.

If $\lambda = 3$ and $\mu = 10$, consistent.

If $\lambda \neq 3$, consistent and the solution is unique.

9.5 Cayley's Hamilton Theorem

Definition: 9.5.1

An expression of the form $A_0 + A_1x + A_2x^2 + \dots + A_nx^n$ where $A_0, A_1, \dots, A_n$ are square matrices of the same order and $A_n \neq 0$ is called a **matrix polynomial** of degree n .

For example, $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} + \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}x + \begin{pmatrix} 2 & 0 \\ 3 & 1 \end{pmatrix}x^2$ is a matrix polynomial of degree 2 and it is simply the matrix $\begin{pmatrix} 1+x+2x^2 & 2+x \\ 2x+3x^2 & 3+x+x^2 \end{pmatrix}$.

Definition: 9.5.2

Let A be any square matrix of order n and let I be the identity matrix of order n . then the matrix polynomial given by $A - xI$ is called the **characteristic matrix** of A .

The determinant $|A - xI|$ which is an ordinary polynomial in x of degree n is called the **characteristic polynomial** of A .

The equation $|A - xI| = 0$ is called the **characteristic equation** of A .

Example: 1

$$\text{Let } A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

Then the characteristic matrix of A is $A - xI$ given by

$$\begin{aligned} A - xI &= \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - x \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1-x & 2 \\ 3 & 4-x \end{pmatrix} \end{aligned}$$

$\therefore$ The characteristic polynomial of A is

$$\begin{aligned} |A - xI| &= \begin{vmatrix} 1-x & 2 \\ 3 & 4-x \end{vmatrix} \\ &= (1-x)(4-x) - 6 \\ &= x^2 - 5x - 2 \end{aligned}$$

$\therefore$ The characteristic equation of A is $|A - xI| = 0$

$\therefore x^2 - 5x - 2 = 0$ is the characteristic equation of A.

Example: 2

$$\text{Let } A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix}.$$

The characteristic matrix of A is $A - xI$ given by

$$A - xI = \begin{pmatrix} 1-x & 0 & 2 \\ 0 & 1-x & 2 \\ 1 & 2 & -x \end{pmatrix}$$

$$\text{The characteristic polynomial of A is } |A - xI| = \begin{vmatrix} 1-x & 0 & 2 \\ 0 & 1-x & 2 \\ 1 & 2 & -x \end{vmatrix}$$

$$= (1-x)[(1-x)(-x) - 4] - 2(1-x)]$$

$$= -x(1-x)^2 - 4(1-x) - 2 + 2x$$

$$= -x^3 + 2x^2 - x - 4 + 4x - 2 + 2x$$

$$= -x^3 + 2x^2 + 5x - 6$$

$\therefore$ The characteristic equation of A is

$$-x^3 + 2x^2 + 5x - 6 = 0$$

$$\text{(i.e.) } x^3 - 2x^2 - 5x + 6 = 0$$

Theorem: 9.31 (Cayley's Hamilton Theorem)

Any square matrix A satisfies its characteristic equation.

Space for Hints

(i.e.) if $a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ is the characteristic polynomial of degree n of A then

$$a_0I + a_1A + a_2A^2 + \dots + a_nA^n = 0.$$

Proof:

Let A be a square matrix of order n .

$$\text{Let } |A - xI| = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad (1)$$

be the characteristic polynomial of A .

Now, $\text{adj}(A - xI)$ is a matrix polynomial of degree $n - 1$ since each entry of the matrix $\text{adj}(A - xI)$ is a cofactor of $A - xI$ and hence is a polynomial of degree $\leq n - 1$.

$$\therefore \text{Let } \text{adj}(A - xI) = B_0 + B_1x + B_2x^2 + \dots + B_{n-1}x^{n-1}. \quad (2)$$

Now, $(A - xI)\text{adj}(A - xI) = |A - xI|I$.

$$(\because (\text{adj}A)A = A(\text{adj}A) = |A|I)$$

$$\begin{aligned} \therefore (A - xI)(B_0 + B_1x + \dots + B_{n-1}x^{n-1}) \\ = (a_0 + a_1x + \dots + a_nx^n)I \text{ using (1) and (2).} \end{aligned}$$

$\therefore$ Equating the coefficients of the corresponding powers of x we get

$$AB_0 = a_0I$$

$$AB_1 - B_0 = a_1I$$

$$AB_2 - B_1 = a_2I$$

$$\dots \dots \dots$$

$$\dots \dots \dots$$

$$AB_{n-1} - B_{n-2} = a_{n-1}I$$

$$-B_{n-1} = a_nI$$

Pre-multiplying the above equations by $I, A, A^2, \dots, A^n$ respectively and adding we get

$$a_0I + a_1A + a_2A^2 + \dots + a_nA^n = 0.$$

Note:

The inverse of a non-singular matrix can be calculated by using the Cayley Hamilton theorem as follows.

Let $a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ be the characteristic polynomial of A .

Then by theorem 9.31 we have

$$a_0I + a_1A + a_2A^2 + \dots + a_nA^n = 0. \quad (3)$$

Since $|A - xI| = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ we get $a_0 = |A|$ (by putting $x=0$).

$\therefore a_0 \neq 0$ ($\because A$ is a non singular matrix.)

$$\therefore I = -\frac{1}{a_0} [a_1A + a_2A^2 + \dots + a_nA^n] \text{ (by 3)}$$

$$\therefore A^{-1} = -\frac{1}{a_0} [a_1I + a_2A + \dots + a_nA^{n-1}]$$

Problems:

Problem: 1

Find the characteristic equation of the matrix

$$A = \begin{pmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{pmatrix}$$

Solution:

The characteristic equation of A is given by $|A - \lambda I| = 0$.

$$\text{(i.e.) } \begin{vmatrix} 8-\lambda & -6 & 2 \\ -6 & 7-\lambda & -4 \\ 2 & -4 & 3-\lambda \end{vmatrix} = 0.$$

$$(8-\lambda)[(7-\lambda)(3-\lambda)-16] + 6[-6(3-\lambda)+8] + 2[24-2(7-\lambda)] = 0$$

$$\text{(i.e.) } (8-\lambda)(\lambda^2 - 10\lambda + 5) + 6(6\lambda - 10) + 2(2\lambda + 10) = 10$$

$$\text{(i.e.) } (8\lambda^2 - 80\lambda + 40 - \lambda^3 + 10\lambda^2 - 5\lambda) + (36\lambda - 60) + (4\lambda + 20) = 0$$

$$\text{(i.e.) } \lambda^3 - 18\lambda^2 + 45\lambda = 0, \text{ which represents the characteristic equation of } A.$$

Problem: 2

Show that the non-singular matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$ satisfies the

equation $A^2 - 2A - 5I = 0$. Hence evaluate A^{-1} .

Space for Hints**Solution:**

The characteristic polynomial of A is

$$|A - xI| = \begin{vmatrix} 1-x & 2 \\ 3 & 1-x \end{vmatrix} = x^2 - 2x - 5.$$

$\therefore$ By Cayley – Hamilton theorem $A^2 - 2A - 5I = 0$.

$$\therefore I = \frac{1}{5}(A^2 - 2A).$$

$$\therefore A^{-1} = \frac{1}{5}(A - 2I)$$

$$= \frac{1}{5} \left[\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right]$$

$$= \frac{1}{5} \begin{pmatrix} -1 & 2 \\ 3 & -1 \end{pmatrix}.$$

Problem: 3

Show that the matrix $A = \begin{pmatrix} 2 & -3 & 1 \\ 3 & 1 & 3 \\ -5 & 2 & -4 \end{pmatrix}$ satisfies the equation

$$A(A - I)(A + 2I) = 0.$$

Solution:

The characteristic polynomial of A is

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & -3 & 1 \\ 3 & 1-\lambda & 3 \\ -5 & 2 & -4-\lambda \end{vmatrix}$$

$$= (2 - \lambda)[(1 - \lambda)(-4 - \lambda) - 6] + 3[3(-4 - \lambda) + 5(1 - \lambda) + 1[6 + 5(1 - \lambda)]]$$

$$= (2 - \lambda)[-4 - \lambda + 4\lambda + \lambda^2 - 6] + 3[-12 - 3\lambda + 5 - 5\lambda] + 6 + 5 - \lambda$$

$$= (2 - \lambda)[\lambda^2 + 3\lambda - 10] + 3[-8\lambda - 7] + 11 - \lambda$$

$$= -\lambda^3 - \lambda^2 + 2\lambda \text{ (after simplification).}$$

$\therefore$ By Cayley – Hamilton theorem $-A^3 - A^2 + 2A = 0$.

(i.e.) $A^3 + A^2 - 2A = 0$. Hence $A(A^2 + A - 2I) = 0$.

$$\therefore A(A + 2I)(A - I) = 0.$$

Problem: 4

Using Cayley – Hamilton theorem find the inverse of the matrix

$$A = \begin{pmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{pmatrix}$$

Solution:

$$\text{Let } A = \begin{pmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{pmatrix}.$$

The characteristic polynomial of $A = |A - xI|$

$$= \begin{vmatrix} 7-x & 2 & -2 \\ -6 & -1-x & 2 \\ 6 & 2 & -1-x \end{vmatrix}$$

$$= (7-x)[(1+x)^2 - 4] - 2[6(1+x) - 12] - 2[-12 + 6(1+x)]$$

$$= (7-x)(x^2 + 2x - 3) - 12(x-1) - 12(x-1)$$

$$= 7x^2 + 14x - 21 - x^3 - 2x^2 + 3x - 12x + 12 - 12x + 12$$

$$= -x^3 + 5x^2 - 7x + 3.$$

$\therefore$ By Cayley – Hamilton theorem

$$-A^3 + 5A^2 - 7A + 3I_3 = 0.$$

$$\therefore A^3 - 5A^2 + 7A - 3I_3 = 0.$$

$$\therefore 3I_3 = A^3 - 5A^2 + 7A.$$

$$\therefore I_3 = \frac{1}{3}(A^3 - 5A^2 + 7A).$$

Pre (or post) multiplying by A^{-1} on both sides we get

$$A^{-1} = \frac{1}{3}[A^2 - 5A + 7I_3] \quad (1)$$

$$\begin{aligned} \text{Now, } A^2 &= \begin{pmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{pmatrix} \begin{pmatrix} 7 & 2 & -2 \\ -6 & -1 & 2 \\ 6 & 2 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 25 & 8 & -8 \\ -24 & -7 & 8 \\ 24 & 8 & -7 \end{pmatrix} \quad (\text{verify}). \end{aligned}$$

$\therefore$ From (1)

Space for Hints

$$A^{-1} = \frac{1}{3} \left[\begin{pmatrix} 25 & 8 & -8 \\ -24 & -7 & 8 \\ 24 & 8 & -7 \end{pmatrix} - \begin{pmatrix} 35 & 10 & -10 \\ -30 & -5 & 10 \\ 30 & 10 & -5 \end{pmatrix} + \begin{pmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{pmatrix} \right]$$

$$= \frac{1}{3} \begin{pmatrix} -3 & -2 & 2 \\ 6 & 5 & -2 \\ -6 & -2 & 5 \end{pmatrix}$$

Problem: 5

Find the inverse of the matrix $\begin{pmatrix} 3 & 3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{pmatrix}$ using Cayley – Hamilton

theorem.

Solution:

The characteristic polynomial of $A = |A - xI| = \begin{vmatrix} 3-x & 3 & 4 \\ 2 & -3-x & 4 \\ 0 & -1 & 1-x \end{vmatrix}$

$$= -x^3 + x^2 + 11x - 11 \text{ (verify).}$$

$\therefore$ By Cayley – Hamilton theorem

$$-A^3 + A^2 + 11A - 11I_3 = 0.$$

$$\therefore A^3 - A^2 - 11A + 11I_3 = 0.$$

$$\text{Hence } 11I_3 = -(A^3 - A^2 - 11A).$$

$$\therefore I_3 = -\frac{1}{11}[A^3 - A^2 - 11A].$$

Pre (post) multiplying by A^{-1} on both sides we get

$$A^{-1} = -\frac{1}{11}[A^2 - A - 11I_3]$$

$$= -\frac{1}{11} \left[\begin{pmatrix} 15 & -4 & 28 \\ 0 & 11 & 0 \\ -2 & 2 & -3 \end{pmatrix} - \begin{pmatrix} 3 & 3 & 4 \\ 2 & -3 & 4 \\ 0 & -1 & 1 \end{pmatrix} - 11 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right] = \begin{pmatrix} -\frac{1}{11} & \frac{7}{11} & -\frac{24}{11} \\ \frac{2}{11} & -\frac{3}{11} & \frac{4}{11} \\ \frac{11}{11} & -\frac{3}{11} & \frac{15}{11} \end{pmatrix}$$

Problem: 6

Verify Cayley Hamilton's theorem for the matrix $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$.

Solution:

The characteristic equation of A is $|A - \lambda I| = 0$.

$$\therefore \begin{vmatrix} 1-\lambda & 2 \\ 4 & 3-\lambda \end{vmatrix} = 0$$

$$\therefore (1-\lambda)(3-\lambda) - 8 = 0$$

$$\therefore \lambda^2 - 4\lambda - 5 = 0.$$

By Cayley Hamilton's theorem A satisfies its characteristic equation.

$\therefore$ We have $A^2 - 4A - 5I = 0$.

$$\begin{aligned} \text{Now, } A^2 &= \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 9 & 8 \\ 16 & 17 \end{pmatrix} \end{aligned}$$

$$4A = \begin{pmatrix} 4 & 8 \\ 16 & 12 \end{pmatrix} \text{ and } 5I = \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix}$$

$$\therefore A^2 - 4A - 5I = \begin{pmatrix} 9 & 8 \\ 16 & 17 \end{pmatrix} - \begin{pmatrix} 4 & 8 \\ 16 & 12 \end{pmatrix} - \begin{pmatrix} 5 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0$$

Thus Cayley Hamilton's theorem is verified.

Problem: 7

Using Cayley Hamilton's theorem for the matrix

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix}$$

Find (i) A^{-1} (ii) A^4 .

Solution:

The characteristic equation of A is $|A - \lambda I| = 0$.

$$\therefore \begin{vmatrix} 1-\lambda & 0 & -2 \\ 2 & 2-\lambda & 4 \\ 0 & 0 & 2-\lambda \end{vmatrix} = 0$$

$$(i.e) \quad \lambda^3 - 5\lambda^2 + 8\lambda - 4 = 0. \text{ (Verify)}$$

$\therefore$ By Cayley Hamilton's theorem

$$A^3 - 5A^2 + 8A - 4I = 0 \quad (1)$$

Space for Hints

Space for Hints

$$4I = A^3 - 5A^2 + 8A$$

I. To find A^{-1} pre multiplying by A^{-1} we get

$$\begin{aligned} 4A^{-1} &= A^{-1}A^3 - 5A^{-1}A^2 + 8A^{-1}A \\ &= A^2 - 5A + 8I \end{aligned}$$

$$\therefore A^{-1} = \frac{1}{4}[A^2 - 5A + 8I] \quad (2)$$

$$\text{Now, } A^2 = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & -6 \\ 6 & 4 & 12 \\ 0 & 0 & 4 \end{pmatrix}$$

From (2)

$$A^{-1} = \frac{1}{4} \left[\begin{pmatrix} 1 & 0 & -6 \\ 6 & 4 & 12 \\ 0 & 0 & 4 \end{pmatrix} - \begin{pmatrix} 5 & 0 & -10 \\ 10 & 10 & 20 \\ 0 & 0 & 10 \end{pmatrix} + \begin{pmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 8 \end{pmatrix} \right]$$

$$= \frac{1}{4} \begin{pmatrix} 4 & 0 & 4 \\ 4 & 2 & -8 \\ 0 & 0 & 2 \end{pmatrix}$$

$$\therefore A^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & \frac{1}{2} & -2 \\ 0 & 0 & -\frac{1}{2} \end{pmatrix}$$

II. To find A^4 .

$$\text{From (1) } A^3 = 5A^2 - 8A + 4I$$

$$\therefore A^4 = 5A^3 - 8A^2 + 4A$$

$$= 5[5A^2 - 8A + 4I] - 8A^2 + 4A \quad (\text{using (1)})$$

$$= 17A^2 - 36A + 20I$$

$$= 17 \begin{pmatrix} 1 & 0 & -6 \\ 6 & 4 & 12 \\ 0 & 0 & 4 \end{pmatrix} - 36 \begin{pmatrix} 1 & 0 & -2 \\ 2 & 2 & 4 \\ 0 & 0 & 2 \end{pmatrix} + 20 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} 17 & 0 & -102 \\ 102 & 68 & 204 \\ 0 & 0 & 68 \end{pmatrix} - \begin{pmatrix} 36 & 0 & -72 \\ 72 & 72 & 144 \\ 0 & 0 & 72 \end{pmatrix} + \begin{pmatrix} 20 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 20 \end{pmatrix}$$

$$\therefore A^4 = \begin{pmatrix} 1 & 0 & -30 \\ 30 & 16 & 60 \\ 0 & 0 & 16 \end{pmatrix}$$

Exercises:

1. Obtain the characteristic polynomial for the following matrices.

(i). $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

(ii). $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$

2. Find the characteristic equation of the following matrices.

(i) $\begin{pmatrix} -b & -c \\ 1 & 0 \end{pmatrix}$

(ii) $\begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$

(iii) $\begin{pmatrix} 8 & -6 & 2 \\ -6 & 7 & -8 \\ 2 & -4 & 3 \end{pmatrix}$

(iv) $\begin{pmatrix} -b & -c & -d \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$

3. Verify Cayley – Hamilton theorem for the matrix $A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}$.

and hence find A^{-1} .

4. If $A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix}$, prove that $A^3 - 2A^2 - 5A + 6I = 0$.

5. Verify Cayley – Hamilton theorem for A and hence find A^{-1} .

Space for Hints

$$(a) \quad A = \begin{pmatrix} 2 & -1 & 1 \\ -15 & 6 & -5 \\ 5 & -2 & 2 \end{pmatrix}$$

$$(b) \quad A = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 2 & 0 & 3 \end{pmatrix}$$

$$(c) \quad A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}$$

$$(d) \quad A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}$$

6. If $A = \begin{pmatrix} 2 & 4 \\ 1 & 1 \end{pmatrix}$ find A^3 and A^{-3} .

7. Verify that the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 4 \\ 3 & 1 & -1 \end{pmatrix}$ satisfies its own

characteristic equation and hence find A^{-1} and A^4 .

9.6 Eigen Values and Eigen Vectors

Definition: 9.6.1

Let A be an $n \times n$ matrix. A number λ is called an **eigen value** of A

if there exists a non-zero vector $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ such that $AX = \lambda X$ and X is

called an **eigen vector** corresponding to the eigen value λ .

Remark: 1

If X is an eigen vector corresponding to the eigen value λ of A , then αX where α is any non-zero number, is also an eigen vector corresponding to λ .

Remark: 2

Let X be an eigen vector corresponding to the eigen value λ of A . Then $AX = \lambda X$ so that $(A - \lambda I)X = 0$. Thus X is a non-trivial solution of

the system of homogeneous linear equations $(A - \lambda I)X = 0$. Hence $|A - \lambda I| = 0$, which is characteristic polynomial of A.

$$\text{Let } |A - \lambda I| = a_0 \lambda^n + a_1 \lambda^{n-1} + \dots + a_n.$$

The roots of this polynomial give the eigen values of A. Hence eigen values are also **characteristic roots**.

Properties of Eigen Values

Property: 1

Let X be an eigen vector corresponding to the eigen values λ_1 and λ_2 . Then $\lambda_1 = \lambda_2$.

Proof:

By definition $X \neq 0$, $AX = \lambda_1 X$ and $AX = \lambda_2 X$

$$\therefore \lambda_1 X = \lambda_2 X$$

$$\therefore (\lambda_1 - \lambda_2)X = 0$$

Since $X \neq 0$, $\lambda_1 = \lambda_2$.

Property: 2

Let A be a square matrix.

Then (i) the sum of the eigen values of A is equal to the sum of the diagonal elements (trace) of A.

(ii) Product of eigen values of A is $|A|$.

Poof:

$$(i) \quad \text{Let } A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

The eigen values of A are the roots of the characteristic equation

$$|A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0. \quad (1)$$

$$\text{Let } |A - \lambda I| = a_0 \lambda^n + a_1 \lambda^{n-1} + \dots + a_n \quad (2)$$

From (1) and (2) we get

$$a_0 = (-1)^n; a_1 = (-1)^{n-1} (a_{11} + a_{22} + \dots + a_{nn}); \quad (3)$$

Space for Hints

Also by putting $\lambda = 0$ in (2) we get $a_n = |A|$

Now let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigen values of A.

$\therefore \lambda_1, \lambda_2, \dots, \lambda_n$ be the eigen values of A.

$$\therefore \lambda_1 + \lambda_2 + \dots + \lambda_n = -\frac{a_1}{a_0} = a_{11} + a_{22} + \dots + a_{nn} \text{ (using (3))}$$

$\therefore$ Sum of the eigen values = trace of A.

(ii) Product of the eigen values = product of the roots

$$= \lambda_1 \lambda_2, \dots, \lambda_n$$

$$= (-1)^n \frac{a_n}{a_0} = \frac{(-1)^n a_n}{(-1)^n}$$

$$= a_n$$

$$= |A|.$$

Property: 3

The eigen values of A and its transpose A^T are the same.

Proof:

It is enough if we prove that A and A^T have the same characteristic polynomial. Since for any square matrix M, $|M| = |M^T|$ we have,

$$\begin{aligned} |A - \lambda I| &= |(A - \lambda I)^T| = |A^T - (\lambda I)^T| \\ &= |A^T - \lambda I|. \end{aligned}$$

Hence the result.

Property: 4

If λ is an eigen value of a non singular matrix A then $\frac{1}{\lambda}$ is an eigen value of A^{-1} .

Proof:

Let X be an eigen vector corresponding to λ . Then $AX = \lambda X$. Since A is non singular A^{-1} exists.

$$\therefore A^{-1}(AX) = A^{-1}(\lambda X)$$

$$IX = \lambda A^{-1}X$$

$$\therefore A^{-1}X = \left(\frac{1}{\lambda}\right)X.$$

$\therefore \frac{1}{\lambda}$ is an eigen value of A^{-1} .

Corollary:

If $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigen values of a non singular matrix A then $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_n}$ are the eigen values of A^{-1} .

Property: 5

If λ is an eigen value of A then $k\lambda$ is an eigen value of kA where k is a scalar.

Proof:

Let X be an eigen vector corresponding to λ .

$$\text{Then } AX = \lambda X \quad (1)$$

Now,

$$\begin{aligned} (kA)X &= k(AX) \\ &= k(\lambda X) \quad (\text{by (1)}) \\ &= (k\lambda)X. \end{aligned}$$

$\therefore k\lambda$ is an eigen value of kA .

Property: 6

If λ is an eigen value of A then λ^k is an eigen value of A^k where k is any positive integer.

Proof:

Let X be an eigen vector corresponding to λ .

$$\text{Then } AX = \lambda X \quad (1)$$

$$\text{Now, } A^2X = (AA)X = A(AX)$$

$$= A(\lambda X) \quad (\text{by (1)})$$

$$= \lambda(AX)$$

$$= \lambda(\lambda X) \quad (\text{by (1)})$$

$$= \lambda^2 X.$$

$\therefore \lambda^2$ is an eigen value of A^2 .

Proceeding like this we can prove that λ^k is an eigen value of A^k for any positive integer.

Corollary:

If $\lambda_1, \lambda_2, \dots, \lambda_n$ are eigen values of A then $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$ are eigen values of A^k for any positive integer k .

Property: 7

Eigen vectors corresponding to distinct eigen values of a matrix are linearly independent.

Proof:

Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be distinct eigen values of a matrix and let X_i be the eigen vector corresponding to λ_i .

$$\text{Hence } AX_i = \lambda_i X_i \quad (i = 1, 2, \dots, k) \quad (1)$$

Now, suppose $X_1, X_2, \dots, X_k$ are linearly dependent. Then there exist real numbers $\alpha_1, \alpha_2, \dots, \alpha_k$, not all zero, such that $\alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_k X_k = 0$. Among all such relations, we choose one of shortest length, say j .

$$\text{By rearranging the vectors } X_1, X_2, \dots, X_k \text{ we may assume that } \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_j X_j = 0 \quad (2)$$

$$\therefore A(\alpha_1 X_1) + A(\alpha_2 X_2) + \dots + A(\alpha_j X_j) = 0$$

$$\therefore \alpha_1 (AX_1) + \alpha_2 (AX_2) + \dots + \alpha_j (AX_j) = 0$$

$$\therefore \alpha_1 \lambda_1 X_1 + \alpha_2 \lambda_2 X_2 + \dots + \alpha_j \lambda_j X_j = 0 \quad (3)$$

Multiplying (2) by λ_1 and subtracting from (3), we get

$$\alpha_2 (\lambda_1 - \lambda_2) X_2 + \alpha_3 (\lambda_1 - \lambda_3) X_3 + \dots + \alpha_j (\lambda_1 - \lambda_j) X_j = 0 \quad (4)$$

and since $\lambda_1, \lambda_2, \dots, \lambda_j$ are distinct and $\alpha_2, \dots, \alpha_j$ are non-zero we have

$$\alpha_i (\lambda_1 - \lambda_i) \neq 0; \quad i = 2, 3, \dots, j.$$

Thus (4) gives a relation whose length is $j-1$, giving a contradiction.

Hence $X_1, X_2, \dots, X_k$ are linearly independent.

Property: 8

The characteristic roots of a Hermitian matrix are all real.

Hence

$$A = \overline{A}^T \quad (\text{by theorem 9.13}) \quad (1)$$

Let λ be a characteristic root of A and let X be a characteristic vector

corresponding to λ .

$$\therefore AX = \lambda X \quad (2)$$

Now,

$$\begin{aligned} AX = \lambda X &\Rightarrow \bar{X}^T AX = \lambda \bar{X}^T X \\ &\Rightarrow (\bar{X}^T AX)^T = \lambda \bar{X}^T X \quad (\text{since } X^T AX \text{ is a } 1 \times 1 \text{ matrix}) \\ &\Rightarrow X^T A^T (\bar{X}^T)^T = \lambda \bar{X}^T X \\ &\Rightarrow X^T A^T \bar{X} = \lambda \bar{X}^T X \\ &\Rightarrow \overline{X^T A^T \bar{X}} = \overline{\lambda X^T X} \\ &\Rightarrow \bar{X}^T \bar{A}^T X = \bar{\lambda} X^T \bar{X} \\ &\Rightarrow \bar{X}^T AX = \bar{\lambda} X^T \bar{X} \quad (\text{using 1}) \\ &\Rightarrow \bar{X}^T \lambda X = \bar{\lambda} X^T \bar{X} \quad (\text{using 2}) \\ &\Rightarrow \lambda (\bar{X}^T X) = \bar{\lambda} (X^T \bar{X}) \quad (3) \end{aligned}$$

Now,

$$\begin{aligned} \bar{X}^T X &= X^T \bar{X} = \bar{x}_1 x_1 + \bar{x}_2 x_2 + \dots + \bar{x}_n x_n \\ &= |x_1|^2 + |x_2|^2 + \dots + |x_n|^2 \\ &\neq 0 \end{aligned}$$

$\therefore$ From (3) we get $\lambda = \bar{\lambda}$

Hence λ is real.

Corollary:

The characteristic roots of a real symmetric matrix are real.

Proof:

We know that any real symmetric matrix is Hermitian. Hence the result follows from the above property.

Property: 9

The characteristic roots of a skew Hermitian matrix are either purely imaginary or zero.

Proof:

Let A be a skew Hermitian matrix and λ be a characteristic root of A .

$$\therefore |A - \lambda I| = 0$$

$$\therefore |iA - i\lambda I| = 0$$

$\therefore i\lambda$ is a characteristic root of iA .

Since A is skew Hermitian iA is Hermitian (refer results iv theorem 9.14)

$\therefore$ By theorem 9.32 $i\lambda$ is real. Hence λ is purely imaginary or zero.

Corollary:

The characteristic roots of a real skew symmetric matrix are either purely imaginary or zero.

Proof:

We know that any real skew symmetric matrix skew Hermitian.

Hence the result follows from the above property.

Property: 10

Let λ be a characteristic root of an unitary matrix A . Then $|\lambda| = 1$ (i.e) the characteristic roots of a unitary matrix are all the unit modulus.

Proof:

Let λ be a characteristic root of an unitary matrix A and X be a characteristic vector corresponding to λ .

$$\therefore AX = \lambda X \quad (1)$$

Taking conjugate and transpose in (1) we get $(\overline{AX})^T = (\overline{\lambda X})^T$.

$$\therefore \overline{X}^T \overline{A}^T = \overline{\lambda} \overline{X}^T \quad (2)$$

Multiplying (1) and (2) we get

$$(\overline{X}^T \overline{A}^T)(AX) = (\overline{\lambda} \overline{X}^T)(\lambda X)$$

$$\therefore \overline{X}^T (\overline{A}^T A) X = \overline{\lambda} \lambda (\overline{X}^T X)$$

Now, since A is an unitary matrix $\overline{A}^T A = I$.

$$\text{Hence } \overline{X}^T X = (\overline{\lambda} \lambda) \overline{X}^T X$$

Since X is non-zero vector $\overline{X}^T$ is also non-zero vector and

$$\overline{X}^T X = |x_1|^2 + |x_2|^2 + \dots + |x_n|^2 \neq 0 \text{ we get } \lambda \overline{\lambda} = 1.$$

$$\text{Hence } |\lambda|^2 = 1. \text{ Hence } |\lambda| = 1.$$

Corollary:

Let λ be a characteristic root of an orthogonal matrix A . Then $|\lambda| = 1$.

Since any orthogonal matrix is unitary the result follows from property 10.

Property: 11

Zero is an eigen value of A if and only if A is a singular matrix.

Proof:

The eigen values of A are the roots of the characteristic equation

$$|A - \lambda I| = 0. \text{ Now, } 0 \text{ is an eigen value of } A \Leftrightarrow |A - 0I| = 0$$

$$\Leftrightarrow |A| = 0$$

$$\Leftrightarrow A \text{ is a singular matrix}$$

Property: 12

If A and B are two square matrices of the same order then AB and BA have the same eigen values.

Solution:

Let λ be an eigen value of AB and X be an eigen vector corresponding to λ .

$$\therefore (AB)X = \lambda X$$

$$\therefore B(AB)X = B(\lambda X) = \lambda(BX).$$

$$\therefore (BA)(BX) = \lambda(BX)$$

$$\therefore (BA)Y = \lambda Y \text{ where } Y = BX.$$

Hence λ is an eigen value of BA .

Also BX is the corresponding eigen vector.

Property: 13

If P and A are $n \times n$ matrices and P is a nonsingular matrix then A and $P^{-1}AP$ have the same eigen values.

Proof:

$$\text{Let } B = P^{-1}AP.$$

To prove A and B have same eigen values, it is enough to prove that the characteristic polynomials of A and B are the same.

Now,

$$|B - \lambda I| = |P^{-1}AP - \lambda I|$$

$$= |P^{-1}AP - P^{-1}(\lambda I)P|$$

$$= |P^{-1}(A - \lambda I)P|$$

$$= |P^{-1}| |A - \lambda I| |P|$$

$$= |P^{-1}P| |A - \lambda I|$$

$$= |I| |A - \lambda I|$$

$$= |A - \lambda I|.$$

$\therefore$ The characteristic equations of A and $P^{-1}AP$ are the same.

Property: 14

If λ is a characteristic root of A then $f(\lambda)$ is a characteristic root of the matrix $f(A)$ where $f(x)$ is any polynomial.

Proof:

Let $f(x) = a_0x^n + a_1x^{n-1} + \dots + a_{n-1}x + a_n$ where $a_0 \neq 0$ and $a_1, a_2, \dots, a_n$ are all real numbers.

$$\therefore f(A) = a_0A^n + a_1A^{n-1} + \dots + a_{n-1}A + a_nI.$$

Since λ is a characteristic root of A , λ^n is a characteristic root of A^n for any positive integer n (refer Property 6).

$$\therefore A^n X = \lambda^n X$$

$$A^{n-1} X = \lambda^{n-1} X$$

$$\dots\dots\dots$$

$$\dots\dots\dots$$

$$AX = \lambda X$$

$$\therefore a_0 A^n X = a_0 \lambda^n X$$

$$a_1 A^{n-1} X = a_1 \lambda^{n-1} X$$

$$\dots\dots\dots$$

$$\dots\dots\dots$$

$$a_{n-1} AX = a_{n-1} \lambda X$$

Adding the above equations we have

$$a_0 A^n X + a_1 A^{n-1} X + \dots\dots\dots a_{n-1} AX$$

$$= a_0 \lambda^n X + a_1 \lambda^{n-1} X + \dots\dots\dots + a_{n-1} \lambda X$$

$$\therefore (a_0 A^n + a_1 A^{n-1} + \dots\dots\dots + a_{n-1} A) X$$

$$= (a_0 \lambda^n + a_1 \lambda^{n-1} + \dots\dots\dots + a_{n-1} \lambda) X$$

$$\therefore (a_0 A^n + a_1 A^{n-1} + \dots\dots\dots + a_{n-1} A + a_n I) X$$

$$= (a_0\lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n)X$$

$$\therefore f(A)X = f(\lambda)X$$

Hence $f(\lambda)$ is a characteristic root of $f(A)$.

Problems:

Problem: 1

If X_1, X_2 are eigen vectors corresponding to an eigen value λ then $aX_1 + bX_2$ (a, b non-zero scalars) is also an eigen vector corresponding to λ .

Solution:

Since X_1 and X_2 are given vectors corresponding to λ , we have

$$AX_1 = \lambda X_1 \text{ and } AX_2 = \lambda X_2.$$

Hence $A(aX_1) = \lambda(aX_1)$ and $A(bX_2) = \lambda(bX_2)$.

$$\therefore A(aX_1 + bX_2) = \lambda(aX_1 + bX_2).$$

$\therefore aX_1 + bX_2$ is an eigen vector corresponding to λ .

Problem: 2

If the eigen values of $A = \begin{pmatrix} 3 & 10 & 5 \\ -2 & -3 & -4 \\ 3 & 5 & 7 \end{pmatrix}$ are 2, 2, 3 find the

eigen values of A^{-1} and A^2 .

Solution:

Since 0 is not an eigen value of A , A is a non singular matrix and hence A^{-1} exists.

Eigen values of A^{-1} are $\frac{1}{2}, \frac{1}{2}, \frac{1}{3}$ and eigen values of A^2 are

$2^2, 2^2, 3^2$.

Problem: 3

Find the eigen values of A^5 when $A = \begin{pmatrix} 3 & 0 & 0 \\ 5 & 4 & 0 \\ 3 & 6 & 1 \end{pmatrix}$

Solution:

The characteristic equation of A is obviously

$$(3-\lambda)(4-\lambda)(1-\lambda)=0$$

Hence the eigen values of A are 3, 4, 1.

$\therefore$ The eigen values of A^5 are $3^5, 4^5, 1^5$.

Problem: 4

Find the sum and product of the eigen values of the matrix

$$\begin{pmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 1 & -1 & 3 \end{pmatrix} \text{ without actually finding the eigen values.}$$

Solution:

$$\text{Let } A = \begin{pmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 1 & -1 & 3 \end{pmatrix}$$

Sum of the eigen values = trace of $A = 3 + (-2) + 3 = 4$.

Product of the eigen values = $|A|$.

$$\begin{aligned} \text{Now, } |A| &= \begin{vmatrix} 3 & -4 & 4 \\ 1 & -2 & 4 \\ 1 & -1 & 3 \end{vmatrix} \\ &= 3(-6+4) + 4(3-4) - 4(-1+2) \\ &= -6 - 4 - 4 = -14. \end{aligned}$$

$\therefore$ Product of the eigen values = -14.

Problem: 5

Find the characteristic roots of the matrix $\begin{pmatrix} \cos \theta & -\sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$

Solution:

$$\text{Let } A = \begin{pmatrix} \cos \theta & -\sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$$

The characteristic equation of A is given by $|A - \lambda I| = 0$.

$$\therefore \begin{vmatrix} \cos \theta - \lambda & -\sin \theta \\ -\sin \theta & \cos \theta - \lambda \end{vmatrix} = 0.$$

$$\therefore (\cos \theta - \lambda)^2 - \sin^2 \theta = 0.$$

$$\therefore (\cos \theta - \lambda)^2 - \sin^2 \theta = 0.$$

$$\therefore (\cos \theta - \lambda - \sin \theta)(\cos \theta - \lambda + \sin \theta) = 0.$$

$$\therefore [\lambda - (\cos \theta - \sin \theta)][\lambda - (\cos \theta + \sin \theta)] = 0.$$

$\therefore$ The two characteristic roots, (the two eigen values) of the matrix are $(\cos \theta - \sin \theta)$ and $(\cos \theta + \sin \theta)$.

Problem: 6

Find the characteristic roots of the matrix $A = \begin{pmatrix} \cos \theta & -\sin \theta \\ -\sin \theta & -\cos \theta \end{pmatrix}$

Solution:

The characteristic equation of A is given by $|A - \lambda I| = 0$.

$$\text{(i.e.) } \begin{vmatrix} \cos \theta - \lambda & -\sin \theta \\ -\sin \theta & -\cos \theta - \lambda \end{vmatrix} = 0$$

$$\therefore -(\cos^2 \theta - \lambda^2) - \sin^2 \theta = 0.$$

$$\therefore \lambda^2 - (\cos^2 \theta + \sin^2 \theta) = 0.$$

$$\therefore \lambda^2 - 1 = 0.$$

$\therefore$ The characteristic roots are 1 and -1.

Problem: 7

Find the sum and product of the eigen values of the matrix

$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ without finding the roots of the characteristic equation.

Solution:

Sum of the eigen values of A = trace of A = $a_{11} + a_{22}$.

Product of the eigen values of A = $|A| = a_{11}a_{22} - a_{12}a_{21}$.

Problem: 8

Verify the statement that the sum of the elements in the diagonal of a matrix is the sum of the eigen values of the matrix

$$A = \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix}$$

Solution:

The characteristic equation of A is $|A - \lambda I| = 0$.

$$(i.e.) \begin{vmatrix} -2-\lambda & 2 & -3 \\ 2 & 1-\lambda & -6 \\ -1 & -2 & -\lambda \end{vmatrix} = 0.$$

$$(i.e.) (-2-\lambda)[(1-\lambda)(-\lambda)-12] - 2[-2\lambda-6] - 3[-4+(1-\lambda)] = 0.$$

$$(i.e.) (-2-\lambda)(\lambda^2 - \lambda - 12) + 4(\lambda + 3) + 3(\lambda + 3) = 0.$$

$$(i.e.) -2\lambda^2 + 2\lambda + 24 - \lambda^3 + \lambda^2 + 12\lambda + 4\lambda + 12 + 3\lambda + 9 = 0.$$

$$(i.e.) -\lambda^3 - \lambda^2 + 21\lambda + 45 = 0.$$

$$(i.e.) \lambda^3 + \lambda^2 - 21\lambda - 45 = 0.$$

This is a cubic equation in λ and hence it has 3 roots and the three roots are the three eigen values of the matrix.

$$\begin{aligned} \text{The sum of the eigen values} &= -\left(\frac{\text{coefficient of } \lambda^2}{\text{coefficient of } \lambda^3}\right) \\ &= -1. \end{aligned}$$

The sum of the elements on the diagonal of the matrix

$$A = -2 + 1 + 0 = -1.$$

Hence the result.

Problem: 9

$$\text{The product of two eigen values of the matrix } A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix} \text{ is}$$

16. Find the third eigen value. What is the sum of the eigen values of A?

Solution:

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigen values of A.

Given, product of 2 eigen values (say) λ_1, λ_2 is 16.

$$\therefore \lambda_1 \lambda_2 = 16$$

We know that the product of the eigen values is $|A|$.

$$(i.e.) \lambda_1 \lambda_2 \lambda_3 = \begin{vmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{vmatrix}$$

$$\begin{aligned} (i.e.) 16\lambda_3 &= 6(9-1) + 2(-6+2) + 2(2-6) \\ &= 48 - 8 - 8 \end{aligned}$$

$$= 32$$

$$\therefore \lambda_3 = 2$$

$\therefore$ The third eigen value is 2.

Also we know that the sum of the eigen values of A = trace of $A = 6 + 3 + 3 = 12$.

Problem: 10

The product of two eigen values of the matrix $A = \begin{pmatrix} 2 & 2 & -7 \\ 2 & 1 & 2 \\ 0 & 1 & -3 \end{pmatrix}$

is -12. Find the eigen values of A .

Solution:

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigen values of A . Given product of 2 eigen values, say, λ_1 and λ_2 is -12.

$$\therefore \lambda_1 \lambda_2 = -12 \quad (1)$$

We know that the product of the eigen values is $|A|$.

$$\therefore \lambda_1 \lambda_2 \lambda_3 = \begin{vmatrix} 2 & 2 & -7 \\ 2 & 1 & 2 \\ 0 & 1 & -3 \end{vmatrix}$$

$$\text{i.e. } 12\lambda_3 = -12$$

$$\therefore \lambda_3 = -1 \quad (2)$$

Also we know sum of the eigen values = Trace of A .

$$\therefore \lambda_1 + \lambda_2 + \lambda_3 = 2 + 1 - 3 = 0$$

$$\therefore \lambda_1 + \lambda_2 = -1 \quad (\text{using (2)}) \quad (3)$$

Using (3) in (1) we get

$$\lambda_1(-1-\lambda_1) = -12$$

$$\lambda_1^2 + \lambda_1 - 12 = 0$$

$$(\lambda_1 + 4)(\lambda_1 - 3) = 0$$

$$\therefore \lambda_1 = 3 \quad \text{or} \quad -4$$

Putting $\lambda_1 = 3$ in (1) we get $\lambda_2 = -4$. Or putting $\lambda_1 = -4$ in (1) we get

$$\lambda_2 = 3$$

Thus the three eigen values are 3, -4, 1.

Problem: 11

Find the sum of the squares of the eigen values of $A = \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix}$

Solution:

Let $\lambda_1, \lambda_2, \lambda_3$ be the eigen values of A. We know that $\lambda_1^2, \lambda_2^2, \lambda_3^2$ are the eigen values of A^2 .

$$\begin{aligned} \therefore A^2 &= \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 3 & 1 & 4 \\ 0 & 2 & 6 \\ 0 & 0 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 9 & 5 & 38 \\ 0 & 4 & 42 \\ 0 & 0 & 25 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \therefore \text{Sum of the eigen values of } A^2 &= \text{Trace of } A^2 \\ &= 9 + 4 + 25 \end{aligned}$$

$$\text{(i.e.) } \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 38$$

$$\therefore \text{Sum of the squares of the eigen values of } A = 38.$$

Problem: 12

Find the eigen values and eigen vectors of the matrix.

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{pmatrix}$$

Solution:

The characteristic equation of A is

$$|A - \lambda I| = 0.$$

$$\therefore \begin{vmatrix} 1-\lambda & 1 & 3 \\ 1 & 5-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{vmatrix} = 0$$

$$\therefore (1-\lambda)[(5-\lambda)(1-\lambda)-1] - [(1-\lambda)-3] + 3[1-3(5-\lambda)] = 0.$$

$$(1-\lambda)(\lambda^2 - 6\lambda + 4) + (\lambda + 2) + 3(3\lambda - 14) = 0.$$

$$\lambda^2 - 6\lambda + 4 - \lambda^3 + 6\lambda^2 - 4\lambda + \lambda + 2 + 9\lambda - 42 = 0.$$

$$\therefore -\lambda^3 + 7\lambda^2 - 36 = 0. \text{ Hence } \lambda^3 - 7\lambda^2 + 36 = 0.$$

$$\therefore (\lambda + 2)(\lambda^2 - 9\lambda + 18) = 0.$$

$$\text{Hence } (\lambda + 2)(\lambda - 6)(\lambda - 3) = 0.$$

$$\therefore \lambda = -2, 3, 6 \text{ are the three eigen values.}$$

Case: (i)

Eigen vector corresponding to $\lambda = -2$.

$$\text{Let } X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \text{ be an eigen vector corresponding to } \lambda = -2.$$

$$\text{Hence } AX = -2X.$$

$$\text{(i.e.) } \begin{pmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ 3 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2x_1 \\ -2x_2 \\ -2x_3 \end{pmatrix}$$

$$\therefore x_1 + x_2 + 3x_3 = -2x_1$$

$$x_1 + 5x_2 + x_3 = -2x_2$$

$$3x_1 + x_2 + x_3 = -2x_3$$

$$\therefore 3x_1 + x_2 + 3x_3 = 0 \quad (1)$$

$$x_1 + 7x_2 + x_3 = 0 \quad (2)$$

$$3x_1 + x_2 + 3x_3 = 0 \quad (3)$$

Clearly this system of three equations reduces to two equations only.

From (1) and (2) we get

$$\therefore x_1 = -2k; x_2 = 0; x_3 = 2k.$$

$\therefore$ It has only one independent solution and can be obtained by giving any value of k say $k = 1$.

$$\therefore (-2, 0, 2) \text{ is an eigen vector corresponding to } \lambda = -2.$$

Case: ii

Eigen vector corresponding to $\lambda = 3$.

Then $AX = 3X$ gives

$$-2x_1 + x_2 + 3x_3 = 0$$

$$x_1 + 2x_2 + x_3 = 0$$

$$3x_1 + x_2 - 2x_3 = 0.$$

Taking the first 2 equation we get

$$\frac{x_1}{-5} = \frac{x_2}{5} = \frac{x_3}{-5} = k(\text{say}).$$

$$\therefore x_1 = -k; x_2 = k; x_3 = -k.$$

Taking $k = 1$ (say) $(-1, 1, -1)$ is an eigen vector corresponding to $\lambda = 3$.

Case: iii

Eigen vector corresponding to $\lambda = 6$.

We have $AX = 6X$.

$$\text{Hence } -5x_1 + x_2 + 3x_3 = 0$$

$$x_1 - x_2 + x_3 = 0$$

$$3x_1 + x_2 - 5x_3 = 0$$

Taking the first two equations we get

$$\frac{x_1}{4} = \frac{x_2}{8} = \frac{x_3}{4} = k.$$

$$\therefore x_1 = k; x_2 = 2k; x_3 = k. \text{ It satisfies the third equation also.}$$

Taking $k = 1$ (say) $(1, 2, 1)$ is an eigen vector corresponding to $\lambda = 6$.

Problem: 13

Find the eigen values and eigen vectors of the matrix

$$A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 3 & -1 \\ 2 & -1 & 3 \end{pmatrix}$$

Solution:

The characteristic equation of A is $|A - \lambda I| = 0$.

$$\therefore \begin{vmatrix} 6-\lambda & -2 & 2 \\ -2 & 3-\lambda & -1 \\ 2 & -1 & 3-\lambda \end{vmatrix} = 0.$$

$$\therefore (6-\lambda)[(3-\lambda)^2 - 1] + 2[(2\lambda - 6) + 2] + 2(2 - 6 + 2\lambda) = 0.$$

$$\therefore (6-\lambda)(8 + \lambda^2 - 6\lambda) + 4\lambda - 8 + 4\lambda - 8 = 0.$$

$$\therefore 48 + 6\lambda^2 - 36\lambda - 8\lambda - \lambda^3 + 6\lambda^2 + 8\lambda - 16 = 0.$$

$$\therefore -\lambda^3 + 12\lambda^2 - 36\lambda + 32 = 0.$$

$$\text{Hence } \lambda^3 - 12\lambda^2 + 36\lambda - 32 = 0.$$

$$\therefore (\lambda - 2)(\lambda - 2)(\lambda - 8) = 0.$$

$\therefore$ The eigen values are 2, 2, 8.

We now find the eigen vectors.

Case: i

$$\lambda = 2.$$

The eigen vector $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ is got from $AX = 2X$.

$$\therefore 6x_1 - 2x_2 + 2x_3 = 2x_1$$

$$-2x_1 + 3x_2 - x_3 = 2x_2$$

$$2x_1 - x_2 + 3x_3 = 2x_3$$

$$\therefore 4x_1 - 2x_2 + 2x_3 = 0$$

$$-2x_1 + x_2 - x_3 = 0$$

$$2x_1 - x_2 + x_3 = 0.$$

The above three equations are equivalent to the single equation

$$2x_1 - x_2 + x_3 = 0.$$

The independent eigen vectors can be obtained by giving arbitrary values to any two of the unknowns x_1, x_2, x_3 .

Giving $x_1 = 1; x_2 = 2$ we get $x_3 = 0$.

Giving $x_1 = 3; x_2 = 4$ we get $x_3 = -2$.

$\therefore$ Two independent vectors corresponding to $\lambda = 2$ are $(1, 2, 0)$ and $(3, 4, -2)$.

Case: ii

$$\lambda = 8.$$

The eigen vector $X = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ is got from $AX = 8X$.

$$\therefore -2x_1 - 2x_2 + 2x_3 = 0 \quad (1)$$

$$-2x_1 - 5x_2 - x_3 = 0 \quad (2)$$

$$2x_1 - x_2 - 5x_3 = 0 \quad (3)$$

From (1) and (2) we get

$$\frac{x_1}{12} = \frac{x_2}{-6} = \frac{x_3}{6} = k \text{ (say).}$$

$$\therefore x_1 = 2k; x_2 = -k; x_3 = k.$$

Giving $k = 1$ we get an eigen vector corresponding to 8 as $(2, -1, 1)$.

Problem: 14

Find the eigen values and eigen vectors of the matrix

$$A = \begin{pmatrix} 2 & -2 & 2 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix}$$

Solution:

The characteristic equation of A is $|A - \lambda I| = 0$.

$$\text{(i.e.) } \begin{vmatrix} 2-\lambda & -2 & 2 \\ 1 & 1-\lambda & 1 \\ 1 & 3 & -1-\lambda \end{vmatrix} = 0$$

$$\therefore (2-\lambda)[-(1-\lambda)(1+\lambda)-3] + 2[-(1+\lambda)-1] + 2[3-(1-\lambda)] = 0$$

$$\therefore (2-\lambda)(\lambda^2 - 4) - 2(2+\lambda) + 2(2+\lambda) = 0$$

$$\therefore 2\lambda^2 - 8\lambda^3 - \lambda^3 + 4\lambda - 4 - 2\lambda + 4 + 2\lambda = 0$$

$$\therefore -\lambda^3 + 2\lambda^2 + 4\lambda - 8 = 0$$

$$\text{Hence } \lambda^3 - 2\lambda^2 - 4\lambda + 8 = 0$$

$$\therefore (\lambda - 2)(\lambda^2 - 4) = 0$$

$$\text{Hence } (\lambda - 2)(\lambda - 2)(\lambda + 2) = 0$$

$$\therefore \lambda = 2, 2, -2 \text{ are the three eigen values.}$$

Case: i

$$\lambda = 2.$$

Let $X = (x_1, x_2, x_3)$ be an eigen vector corresponding to $\lambda = 2$, X is got from $AX = 2X$.

$$\text{(i.e.) } \begin{pmatrix} 2 & -2 & 2 \\ 1 & 1 & 1 \\ 1 & 3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \end{pmatrix}$$

$\therefore$ The eigen vector corresponding to $\lambda = 2$ is given by the equations

$$2x_1 - 2x_2 + 2x_3 = 2x_1$$

$$x_1 + x_2 + x_3 = 2x_2$$

$$x_1 + 3x_2 - x_3 = 2x_3$$

$$\text{(i.e.) } -x_2 + x_3 = 0$$

$$x_1 - x_2 + x_3 = 0$$

$$x_1 + 3x_2 - 3x_3 = 0$$

Taking (1) and (2) we get $\frac{x_1}{0} = \frac{x_2}{1} = \frac{x_3}{1} = k$ (say).

$$\therefore x_1 = 0; x_2 = k; x_3 = k.$$

Taking $k = 1$, we get $(0, 1, 1)$ as an eigen vector corresponding to $\lambda = 2$.

Case: ii

$$\lambda = -2$$

Corresponding to $\lambda = -2$ we have $AX = -2X$.

$$\therefore 2x_1 - 2x_2 - 2x_3 = -2x_1$$

$$x_1 + x_2 + x_3 = -2x_2$$

$$x_1 + 3x_2 - x_3 = -2x_3$$

$$\therefore 2x_1 - x_2 + x_3 = 0$$

$$x_1 + 3x_2 + x_3 = 0$$

$$x_1 + 3x_2 + x_3 = 0$$

$\therefore$ Taking the first two equations we get,

$$\frac{x_1}{-4} = \frac{x_2}{-1} = \frac{x_3}{7} = k \text{ (say).}$$

$$\therefore x_1 = -4k; x_2 = -k; x_3 = 7k.$$

Taking $k = 1$ we get $(-4, -1, 7)$ as an eigen vector corresponding to the eigen value $\lambda = -2$.

Exercises:

1. For each of the following matrices find the characteristic vectors corresponding to each characteristic root.

$$(a) \begin{pmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{pmatrix}$$

$$(b) \begin{pmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{pmatrix}$$

$$(c) \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

$$(d) \begin{pmatrix} 1 & 0 & 0 & 0 \\ -1 & i & 0 & 0 \\ 2 & 1/2 & -i & 0 \\ 1/3 & -i & \pi & -1 \end{pmatrix}$$

2. For what value of k is 3 a characteristic root of $\begin{pmatrix} 3 & 1 & -1 \\ 3 & 5 & -k \\ 3 & k & -1 \end{pmatrix}$

3. Find the characteristic roots and the corresponding characteristic

vectors of $A^3 + A^2 + A + I$ if $A = \begin{pmatrix} 1 & -1 & -1 \\ 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$

4. Prove that if λ is a characteristic root of the matrix A then $\lambda - k$ is a characteristic root of $A - kI$.

5. Show that the characteristic root of a triangular matrix are just the diagonal elements of the matrix.

6. Show that every orthogonal matrix of odd order should have 1 or -1 as a characteristic root.

7. Show that if λ is a characteristic root of an unitary matrix then so is

$$\frac{1}{\lambda}$$

Answers:

1.

(a) 0, 3, 15. The corresponding characteristic vectors are (1, 2, 2); (2, -1, -2); (2, -2, 1).

(b) 1, 1, 5; The characteristic vectors corresponding to 1 is a linear combination of $(2, -1, 0)$ and $(1, 0, -1)$. A

characteristic vector corresponding to 5 is $(1, 1, 1)$.

(c) 1, 1, 1. A characteristic vector is $(1, 0, 0)$.

$k = 2$.

BILINEAR FORMS

10.0 Introduction

Consider a finite dimensional, inner product space V over the field $\mathbb{R}$ of real numbers. The inner product is a function from $V \times V$ to $\mathbb{R}$ satisfying the following conditions

$$(i) \quad \langle \alpha u_1 + \beta u_2, v \rangle = \alpha \langle u_1, v \rangle + \beta \langle u_2, v \rangle$$

$$(ii) \quad \langle u, \alpha v_1 + \beta v_2 \rangle = \alpha \langle u, v_1 \rangle + \beta \langle u, v_2 \rangle$$

In other words the inner product is a scalar valued function of the two variables u and v and is a linear function in each of the two variables. This type of scalar valued functions is called **bilinear forms**. In this chapter we study bilinear forms on finite dimensional vector spaces.

10.1 Bilinear forms

Definition: 10.1.1

Let V be a vector space over a field F . A bilinear form on V is a function $f : V \times V \rightarrow F$ such that

$$(i) \quad f(\alpha u_1 + \beta u_2, v) = \alpha f(u_1, v) + \beta f(u_2, v)$$

$$(ii) \quad f(u, \alpha v_1 + \beta v_2) = \alpha f(u, v_1) + \beta f(u, v_2) \quad \text{where} \quad \alpha, \beta \in F \quad \text{and}$$

$$u_1, u_2, v_1, v_2 \in V.$$

In other words f is linear as a function of any one of the two variables when the other is fixed.

Examples:

1. Let V be a vector space over $\mathbb{R}$. Then an inner product on V is a bilinear form on V .

2. Let V be any vector space over a field F . Then the zero function

$\hat{0} : V \times V \rightarrow F$ given by $\hat{0}(u, v) = 0$ is a bilinear form. For,

$$\hat{0}(\alpha u_1 + \beta u_2, v) = 0$$

$$= \alpha 0 + \beta 0$$

$$= \alpha \hat{0}(u_1, v) + \beta \hat{0}(u_2, v)$$

Similarly

$$\hat{0}(u, \alpha v_1 + \beta v_2) = \alpha \hat{0}(u, v_1) + \beta \hat{0}(u, v_2)$$

3. Suppose V is a vector space over a field F . Let f_1 and f_2 be two linear functionals on V . (ie) f_1 and f_2 are linear transformations from V to F . Then $f : V \times V \rightarrow F$ defined by $f(u, v) = f_1(u)f_2(v)$ is a bilinear form.

For,

$$\begin{aligned} f(\alpha u_1 + \beta u_2, v) &= f_1(\alpha u_1 + \beta u_2)f_2(v) \\ &= [\alpha f_1(u_1) + \beta f_1(u_2)]f_2(v) \text{ (since } f_1 \text{ is linear)} \\ &= \alpha f_1(u_1)f_2(v) + \beta f_1(u_2)f_2(v) \\ &= \alpha f(u_1, v) + \beta f(u_2, v). \end{aligned}$$

Similarly,

$$f(u, \alpha v_1 + \beta v_2) = \alpha f(u, v_1) + \beta f(u, v_2).$$

Exercises:

1. Show that the function f defined by

$$f(x, y) = x_1y_1 + x_2y_2 + \dots + x_ny_n \text{ where } x = (x_1, x_2, \dots, x_n)$$

and $y = (y_1, y_2, \dots, y_n)$ is a bilinear form on $V_n(F)$.

2. Which of the following are bilinear forms on $V_2(R)$?

Let $x = (x_1, x_2)$ and $y = (y_1, y_2)$.

- (a) $f(x, y) = 1$.
- (b) $f(x, y) = (x_1 - y_1)^2 + x_2y_2$.
- (c) $f(x, y) = (x_1 + y_1)^2 - (x_1 - y_1)^2$.
- (d) $f(x, y) = x_1y_2 - x_2y_1$.

Answers:

(c) and (d) are bilinear forms.

Notation:

Let V be a vector space over a field F . Then the set of all bilinear forms on V is denoted by $L(V, V, F)$.

Theorem: 10.1

Let V be a vector space over a field F . Then $L(V, V, F)$ is a vector space over F under addition and scalar multiplication defined by

$$(f + g)(u, v) = f(u, v) + g(u, v) \quad \text{and}$$

$$(\alpha f)(u, v) = \alpha f(u, v).$$

Proof:

Let $f, g \in L(V, V, F)$ and $\alpha_1 \in F$.

We claim that $f + g$ and $\alpha_1 f \in L(V, V, F)$.

$$\begin{aligned} (f + g)(\alpha u_1 + \beta u_2, v) &= f(\alpha u_1 + \beta u_2, v) + g(\alpha u_1 + \beta u_2, v) \\ &= \alpha f(u_1, v) + \beta f(u_2, v) + \alpha g(u_1, v) + \beta g(u_2, v) \\ &= \alpha[f(u_1, v) + g(u_1, v)] + \beta[f(u_2, v) + g(u_2, v)] \\ &= \alpha[(f + g)(u_1, v)] + \beta[(f + g)(u_2, v)]. \end{aligned}$$

Similarly we can prove that

$$(f + g)(u, \alpha v_1 + \beta v_2) = \alpha[(f + g)(u, v_1)] + \beta[(f + g)(u, v_2)]$$

Hence $(f + g) \in L(V, V, F)$.

Also

$$\begin{aligned} (\alpha_1 f)(\alpha u_1 + \beta u_2, v) &= \alpha_1 f(\alpha u_1 + \beta u_2, v) \\ &= \alpha_1 [\alpha f(u_1, v) + \beta f(u_2, v)] \\ &= \alpha_1 \alpha f(u_1, v) + \alpha_1 \beta f(u_2, v) \\ &= \alpha[(\alpha_1 f)(u_1, v)] + \beta[(\alpha_1 f)(u_2, v)] \end{aligned}$$

Similarly

$$(\alpha_1 f)(u, \alpha v_1 + \beta v_2) = \alpha[(\alpha_1 f)(u, v_1)] + \beta[(\alpha_1 f)(u, v_2)]$$

$$\therefore \alpha_1 f \in L(V, V, F).$$

The remaining axioms of a vector space can be easily verified.

10.2 Matrix of a bilinear form:

Let f be a bilinear form on V . Fix a basis $\{v_1, v_2, \dots, v_n\}$ for V .

Let $u = \alpha_1 v_1 + \dots + \alpha_n v_n$ and $v = \beta_1 v_1 + \dots + \beta_n v_n$.

Then $f(u, v)$

$$\begin{aligned} &= f(\alpha_1 v_1 + \dots + \alpha_n v_n, \beta_1 v_1 + \dots + \beta_n v_n) \\ &= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \beta_j f(v_i, v_j) \end{aligned}$$

$$= \sum_{i=1}^n \sum_{j=1}^n \alpha_{ij} \alpha_i \beta_j \text{ where } f(v_i, v_j) = a_{ij}$$

$$= (\alpha_1, \dots, \alpha_n) \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \dots \\ \beta_n \end{pmatrix}$$

$\therefore f(u, v) = XAY^T$ where

$X = (\alpha_1, \dots, \alpha_n), A = (a_{ij})$ and $Y = (\beta_1, \dots, \beta_n)$.

The $n \times n$ matrix A is called the **matrix of the bilinear form** with respect to the chosen basis.

Conversely, given any $n \times n$ matrix $A = (a_{ij})$ the $f: V \times V \rightarrow F$ defined by $f(u, v) = XAY^T$ is a bilinear form on V and $f(v_i, v_j) = a_{ij}$.

Also if g is any other bilinear form on V such that $g(v_i, v_j) = a_{ij}$, then $f = g$ (verify).

Problems:

Problem: 1

Let f be the bilinear form defined on $V_2(R)$ by $f(x, y) = x_1 y_1 + x_2 y_2$ where $x = (x_1, x_2)$ and $y = (y_1, y_2)$. Find the matrix of f .

(i) w.r.t. the standard basis $\{e_1, e_2\}$.

(ii) w.r.t. the basis $\{(1,1), (1,2)\}$.

Solution:

$$\begin{aligned} \text{(i)} \quad f(e_1, e_1) &= f((1,0), (1,0)) \\ &= 1 \times 1 + 0 \times 0 = 1 \end{aligned}$$

Similarly

$$f(e_1, e_2) = 0$$

$$f(e_2, e_1) = 0$$

$$f(e_2, e_2) = 1.$$

$$\therefore \text{The matrix of } f \text{ is } \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(ii) Let $v_1 = (1,1)$ and $v_2 = (1,2)$.

Then

$$f(v_1, v_1) = 1 + 1 = 2$$

$$f(v_1, v_2) = 1 + 2 = 3$$

$$f(v_2, v_1) = 1 + 2 = 3$$

$$f(v_2, v_2) = 1 + 4 = 5.$$

$\therefore$ The matrix of f is $\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}$.

Exercises:

- Find the matrix of the bilinear form $f(x, y) = x_1 y_2 - x_2 y_1$ with respect to the standard basis in $V_2(\mathbb{R})$.
- Find the matrix of the bilinear form f defined on $V_3(\mathbb{R})$ by $f(x, y) = x_1 y_1 + x_3 y_3$ where $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ w.r.t.
 - Standard basis
 - $\{(1, 1, 0), (0, 1, 1), (1, 0, 1)\}$.

Answers:

1. $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$

2.

(a) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

(b) $\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 2 \end{pmatrix}$

Theorem: 10.2

Let V be a vector space of dimension n over a field F . Fix a basis $\{v_1, v_2, \dots, v_n\}$ for V . Then the function $\varphi: L(V, V, F) \rightarrow M_n(F)$ which associates with each bilinear form $f \in L(V, V, F)$ the $n \times n$ matrix (a_{ij}) where $f(v_i, v_j) = a_{ij}$ is an isomorphism.

Proof:

Clearly φ is 1 - 1 and onto.

Now, let $f, g \in L(V, V, F)$ and $\alpha \in F$.

Let $\varphi(f) = (a_{ij})$ and $\varphi(g) = (b_{ij})$.

Then $(f + g)(v_i, v_j) = f(v_i, v_j) + g(v_i, v_j)$

$$= a_{ij} + b_{ij}$$

$$\therefore \varphi(f + g) = (a_{ij} + b_{ij}) = (a_{ij}) + (b_{ij})$$

$$= \varphi(f) + \varphi(g).$$

Also $(\alpha f)(v_i, v_j) = \alpha f(v_i, v_j) = \alpha a_{ij}$

$$\therefore \varphi(\alpha f) = (\alpha a_{ij}) = \alpha(a_{ij}) = \alpha\varphi(f).$$

Thus φ is an isomorphism.

Corollary:

$L(V, V, F)$ is a vector space of dimension n^2 .

10.3 Quadratic forms

Definition: 10.3.1

A bilinear form f defined on a vector space V is called a **symmetric bilinear form** if $f(u, v) = f(v, u)$ for all $u, v \in V$.

Examples:

- (i) Let V be a vector space over $\mathbb{R}$. Then any inner product defined on V is a symmetric bilinear form.
- (ii) The bilinear form $\hat{0}$ defined in example 2 of 10.1 is a symmetric bilinear form.
- (iii) Let f be a bilinear form of V . Then the bilinear form f_1 defined by $f_1(u, v) = f(u, v) + f(v, u)$ is a symmetric bilinear form.

Theorem: 10.3

A bilinear form f defined on V is symmetric iff its matrix (a_{ij}) w.r.t any one basis $\{v_1, v_2, \dots, v_n\}$ is symmetric.

Proof:

Let f be a symmetric bilinear form.

Now,

$$a_{ij} = f(v_i, v_j)$$

$$= f(v_j, v_i) \quad (\text{since } f \text{ is symmetric})$$

$$= a_{ji}$$

$\therefore (a_{ij})$ is a symmetric matrix.

Conversely, let (a_{ij}) be a symmetric matrix.

$$\text{Hence } A = A^T \quad (\text{by theorem 9.5})$$

Then

$$f(u, v) = XAY^T$$

$$= (XAY^T)^T \quad (\text{since } XAY^T \text{ is a } 1 \times 1 \text{ matrix})$$

$$= YA^T X^T$$

$$= YAX^T$$

$$= f(v, u)$$

$\therefore f$ is a symmetric bilinear form.

Definition: 10.3.2

Let f be a symmetric bilinear form defined by V . Then the **quadratic form** associated with f is the mapping $q: V \rightarrow F$ defined by $q(v) = f(v, v)$.

The matrix of the bilinear form f is called the matrix of the associated quadratic form q .

Examples:

1. Consider the bilinear form f defined on $V_n(F)$ by

$$f(u, v) = x_1 y_1 + x_2 y_2 + \dots + x_n y_n; u = (x_1, \dots, x_n), v = (y_1, \dots, y_n).$$

Then the quadratic form q associated with f is given by

$$q(u) = f(u, u) = x_1^2 + \dots + x_n^2.$$

2. Let A be a symmetric matrix of order n associated with symmetric bilinear form f . Then the corresponding quadratic form is given by

$$q(X) = XAX^T = \sum_{i,j=1}^n a_{ij} x_i x_j$$

For example, consider the symmetric matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 7 & 6 \end{pmatrix}$$

The quadratic form q determined by A w.r.t. the standard basis for $V_3(R)$ is

given by

$$q(v) = (x_1, x_2, x_3) \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 7 \\ 3 & 7 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= x_1^2 + 4x_2^2 + 6x_3^2 + 4x_1x_2 + 14x_2x_3 + 6x_1x_3.$$

3. Consider the diagonal matrix

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

The quadratic form q determined by A w.r.t. the standard basis for $V_3(\mathbb{R})$ is given by

$$q(v) = (x_1, x_2, x_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$= x_1^2 + 2x_2^2 + 3x_3^2.$$

We say this quadratic form q is in the diagonal form.

4. Consider the quadratic form defined on $V_2(\mathbb{R})$ by

$q(x_1, x_2) = 2x_1^2 + x_1x_2 + x_2^2$. Then the symmetric matrix associated with q can be found as follows.

Let

$$2x_1^2 + x_1x_2 + x_2^2 = (x_1, x_2) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= ax_1^2 + 2bx_1x_2 + cx_2^2.$$

$$\therefore a = 2; b = \frac{1}{2}; c = 1.$$

$$\therefore A = \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}$$

Exercises:

- Find the quadratic forms associated with the following matrices w.r.t. the standard basis.

$$(a) \begin{pmatrix} 2 & -3 & 1 \\ -3 & 2 & 4 \\ 1 & 4 & -5 \end{pmatrix}$$

$$(b) \begin{pmatrix} 1 & 2 & 3 \\ 2 & -2 & -4 \\ 3 & -4 & -3 \end{pmatrix}$$

2. Find the matrices for the following quadratic forms.

$$(a) x_1^2 + 4x_1x_2 + 3x_2^2 \text{ in } V_2(\mathbb{R})$$

$$(b) 2x_1^2 + x_2^2 + 3x_1x_2 \text{ in } V_2(\mathbb{R})$$

$$(c) 2x_1^2 + x_3^2 - 6x_1x_2 \text{ in } V_3(\mathbb{R})$$

$$(d) x_1x_2 \text{ in } V_4(\mathbb{R})$$

$$(e) x_1x_2 + x_4^2 \text{ in } V_4(\mathbb{R}).$$

Answers:

1.

$$(a) 2x_1^2 + 2x_2^2 - 5x_3^2 - 6x_1x_2 + 8x_2x_3 + 2x_1x_3.$$

$$(b) x_1^2 - 2x_2^2 + 3x_3^2 + 4x_1x_2 - 8x_2x_3 + 6x_1x_3.$$

2.

$$(a) \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$

$$(b) \begin{pmatrix} 2 & \frac{3}{2} \\ \frac{3}{2} & 1 \end{pmatrix}$$

$$(c) \begin{pmatrix} 2 & -3 & 0 \\ -3 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$(d) \begin{pmatrix} 0 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(e) \begin{pmatrix} 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Theorem: 10.4

Let f be a symmetric bilinear form defined on V . Let q be the associated quadratic form.

$$(i) \quad f(u, v) = \frac{1}{4} \{q(u+v) - q(u-v)\}$$

$$(ii) \quad f(u, v) = \frac{1}{2} \{q(u+v) - q(u) - q(v)\}$$

Proof:

$$(i) \quad \text{Now } \frac{1}{4} \{q(u+v) - q(u-v)\}$$

$$= \frac{1}{4} \{f(u+v, u+v) - f(u-v, u-v)\}$$

$$= \frac{1}{4} \{f(u, u) + f(u, v) + f(v, u) + f(v, v) - f(u, u) + f(u, v) + f(v, u) - f(v, v)\}$$

$$= \frac{1}{4} \{4f(u, v)\}$$

$$= f(u, v).$$

$$\therefore f(u, v) = \frac{1}{4} \{q(u+v) - q(u-v)\}.$$

$$(ii) \quad \text{Now}$$

$$q(u+v) - q(u) - q(v) = f(u+v, u+v) - f(u, u) - f(v, v)$$

$$= f(u, u) + f(u, v) + f(v, u) + f(v, v) - f(u, u) - f(v, v)$$

$$= f(u, v) + f(v, u)$$

$$= f(u, v) + f(u, v) [\because f \text{ is symmetric}]$$

$$= 2f(u, v)$$

$$\therefore f(u, v) = \frac{1}{2} [q(u+v) - q(u) - q(v)].$$

Note:

The above theorem shows that if f is symmetric bilinear form and q the associated quadratic form, then $f(u, v)$ can be determined from q .

Exercises:

1. If q is a quadratic form prove that
 $q(u + v + w) - q(u + v) - q(v + w) - q(u + w) + q(u) + q(v) + q(w) = 0$.
2. Show that if q_1 is the quadratic form associated with the bilinear form f_1 and q_2 is the quadratic form associated with the bilinear form f_2 then $q_1 + q_2$ is the quadratic form associated with the bilinear form $f_1 + f_2$.

10.4 Reduction to Quadratic forms

In example 3 of the quadratic form in section 10.3 we have seen that a quadratic form associated with a diagonal matrix of order n is of the form

$$a_1 x_1^2 + a_2 x_2^2 + \dots + a_n x_n^2$$

Which is known as the **diagonal form**. Now, we prove that any quadratic form can be reduced to the diagonal form by means of a non-singular linear transformation. The method of reduction which we describe below is due to Lagrange.

Consider the quadratic form

$$\begin{aligned} \varphi &= \varphi(x_1, x_2, \dots, x_n) = \sum_{i,j=1}^n a_{ij} x_i x_j \\ &= a_{11} x_1^2 + \dots + a_{nn} x_n^2 + 2a_{12} x_1 x_2 + \dots + 2a_{n(n-1)} x_n x_{n-1}. \end{aligned}$$

Case: (i)

Suppose at least one of $a_{11}, \dots, a_{nn}$ is not zero. We assume, without loss of generality, that $a_{11} \neq 0$.

Then

$$\begin{aligned} \varphi &= (a_{11} x_1^2 + 2a_{12} x_1 x_2 + \dots + 2a_{1n} x_1 x_n) + \sum_{i,j=2}^n a_{ij} x_i x_j \\ &= a_{11} \left(x_1^2 + 2 \frac{a_{12}}{a_{11}} x_1 x_2 + \dots + 2 \frac{a_{1n}}{a_{11}} x_1 x_n \right) + \varphi_1(x_2, \dots, x_n) \text{ (say)} \end{aligned}$$

$$= a_{11} \left(x_1 + \frac{a_{12}}{a_{11}} x_2 + \dots + \frac{a_{1n}}{a_{11}} x_n \right)^2 + \phi_2(x_2, \dots, x_n) \text{ (say)}$$

Now, putting $y_1 = x_1 + \frac{a_{12}}{a_{11}} x_2 + \dots + \frac{a_{1n}}{a_{11}} x_n$, $y_2 = x_2, \dots, y_n = x_n$, ϕ

reduces to

$$\phi = \alpha_1 y_1^2 + \phi_2(y_2, \dots, y_n) \quad (1)$$

Where $\alpha_1 = a_{11}$

Case: ii

Suppose $a_{11} = a_{22} = \dots = a_{nn} = 0$. We still have $a_{ij} \neq 0$ for some i, j such that $i \neq j$.

Without loss of generality we assume that $a_{12} \neq 0$.

Then the non-singular linear transformation

$$x_1 = y_1, x_2 = y_1 + y_2, x_3 = y_3, \dots, x_n = y_n$$

Changes the quadratic form ϕ to another quadratic form in which the term y_1^2 is present.

Now applying the method of case (i) ϕ can be reduced to the form (1). Treating ϕ_2 in the same way we get

$$\phi_2 = \alpha_2 z_2^2 + \phi_3(z_2, \dots, z_n) \text{ so that}$$

$$\phi = \alpha_1 z_1^2 + \alpha_2 z_2^2 + \phi_3(z_2, \dots, z_n).$$

Continuing this process of reduction we obtain ϕ in the form

$$\phi = \alpha_1 w_1^2 + \dots + \alpha_r w_r^2.$$

Problems:

Problem: 1

Reduce the quadratic form

$$x_1^2 + 4x_1x_2 + 4x_1x_3 + 4x_2^2 + 16x_2x_3 + 4x_3^2 \text{ to the diagonal form.}$$

Solution:

$$\text{Let } \phi = x_1^2 + 4x_1x_2 + 4x_1x_3 + 4x_2^2 + 16x_2x_3 + 4x_3^2$$

$$= (x_1 + 2x_2 + 2x_3)^2 + 8x_2x_3$$

Putting $x_1 + 2x_2 + 2x_3 = y_1$, $x_2 = y_2$, $x_3 = y_2 + y_3$ we get

$$= y_1^2 + 8(y_2 + \frac{1}{2}y_3)^2 - 2y_3^2.$$

Putting $z_1 = y_1, z_2 = y_2 + \frac{1}{2}y_3, z_3 = y_3$ we get

$$\phi = z_1^2 + 8z_2^2 - 2z_3^2 \text{ where } z_1 = x_1 + 2x_2 + 2x_3$$

$$z_2 = \frac{1}{2}(x_2 + x_3).$$

$$z_3 = x_3 - x_2.$$

Problem: 2

Reduce the quadratic form $2x_1x_2 - x_1x_3 + x_1x_4 - x_2x_3 + x_2x_4 - 2x_3x_4$ to the diagonal form using Lagrange's method.

Solution:

$$\text{Let } \phi = 2x_1x_2 - x_1x_3 + x_1x_4 - x_2x_3 + x_2x_4 - 2x_3x_4.$$

Putting $x_1 = y_1; x_2 = y_1 + y_2; x_3 = y_3$ and $x_4 = y_4$, we get

$$\begin{aligned} \phi &= 2y_1^2 + 2y_1y_2 - 2y_1y_3 + 2y_1y_4 - y_2y_3 + y_2y_4 - 2y_3y_4 \\ &= 2(y_1^2 + y_1y_2 - y_1y_3 + y_1y_4) - y_2y_3 + y_2y_4 - 2y_3y_4 \\ &= 2\left(y_1 + \frac{1}{2}y_2 - \frac{1}{2}y_3 + \frac{1}{2}y_4\right)^2 - \frac{1}{2}y_2^2 - \frac{1}{2}y_3^2 - \frac{1}{2}y_4^2 - y_3y_4 \end{aligned}$$

Putting $z_1 = y_1 + \frac{1}{2}y_2 - \frac{1}{2}y_3 + \frac{1}{2}y_4; z_2 = y_2; z_3 = y_3; \text{ and } z_4 = y_4$ we get

$$\begin{aligned} \phi &= 2z_1^2 - \frac{1}{2}z_2^2 - \frac{1}{2}z_3^2 - \frac{1}{2}z_4^2 - z_3z_4 \\ &= 2z_1^2 - \frac{1}{2}z_2^2 - \frac{1}{2}(z_3^2 + 2z_3z_4 + z_4^2) \\ &= 2z_1^2 - \frac{1}{2}z_2^2 - \frac{1}{2}(z_3 + z_4)^2 \end{aligned}$$

Putting $w_1 = z_1, w_2 = z_2, w_3 = z_3 + z_4, w_4 = w_4$

$$\text{We get } \phi = 2w_1^2 - \frac{1}{2}w_2^2 - \frac{1}{2}w_3^2$$

Where

$$w_1 = \frac{1}{2}x_1 + \frac{1}{2}x_2 - \frac{1}{2}x_3 + \frac{1}{2}x_4$$

$$w_2 = -x_1 + x_2; w_3 = x_3 + x_4; w_4 = x_4.$$

Exercises:

Reduce the following quadratic forms to diagonal form.

1. $x_1^2 + 2x_2^2 - 7x_3^2 - 4x_1x_2 + 8x_1x_3$

2. $2x_1^2 + 5x_2^2 + 19x_3^2 - 24x_4^2 + 8x_1x_2 + 12x_1x_3 + 8x_1x_4 + 18x_2x_3 - 8x_2x_4 - 16x_3x_4$

3. $2x_1x_2 - x_1x_3 + x_2x_3$

4. $-2x_1x_2 + 2x_2x_3 - 2x_3x_4 + 2x_1x_4$

5. $(x_1 x_2 x_3) \begin{pmatrix} 1 & 2 & 4 \\ 2 & 6 & -2 \\ 4 & -2 & 18 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

6. $(x_1 x_2 x_3) \begin{pmatrix} 0 & 1 & 2 \\ 1 & 1 & -1 \\ 2 & -1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

Answers:

1. $z_1^2 - 2z_2^2 + 9z_3^2$

2. $2z_1^2 - 3z_2^2 + 4z_3^2$

3. $2z_1^2 - \frac{1}{2}z_2^2 + \frac{1}{2}z_3^2$

4. $z_1^2 - z_2^2 + z_3^2 - z_4^2$

5. $z_1^2 + 2z_2^2 - 48z_3^2$

6. $z_1^2 - z_2^2 + 8z_4^2$

